



Représentation d'un environnement par un système multi-capteurs : fusion et interprétation de scène

Bassem Alsahwa

► To cite this version:

Bassem Alsahwa. Représentation d'un environnement par un système multi-capteurs : fusion et interprétation de scène. Vision par ordinateur et reconnaissance de formes [cs.CV]. Télécom Bretagne; Université de Rennes 1, 2014. Français. NNT : . tel-01206283

HAL Id: tel-01206283

<https://hal.science/tel-01206283>

Submitted on 28 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE / Télécom Bretagne

sous le sceau de l'Université européenne de Bretagne

pour obtenir le grade de Docteur de Télécom Bretagne

En accréditation conjointe avec l'Ecole doctorale Matisse

Mention : Traitement du signal et Télécommunications

présentée par

Bassem Alsahwa

préparée dans le département Image et Traitement d'information

Représentation d'un environnement par un système multi-capteurs : fusion et interprétation de scène

Thèse soutenue le 18 décembre 2014

Devant le jury composé de :

Ali Khenchaf

Professeur, ENSTA Bretagne / président

Jean-Paul Haton

Professeur, Université Henri Poincaré, Nancy 1 / rapporteur

Kamel Hamrouni

Professeur, Ecole Nationale d'Ingénieurs de Tunis / rapporteur

El-hadi Zahzah

Maître de conférences, Université de la Rochelle / examinateur

Didier Guériot

Maître de conférences, Télécom Bretagne / examinateur

Basel Solaiman

Professeur, Télécom Bretagne / Directeur de thèse

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Rennes1

Ecole Doctorale – Matisse

REPRÉSENTATION D'UN ENVIRONNEMENT PAR UN SYSTÈME MULTI-CAPTEURS : FUSION ET INTERPRÉTATION DE SCÈNE.

Thèse de Doctorat

Mention : Traitement du Signal et Télécommunication

Présentée par **Bassem Alsahwa**

Département : Image et Traitement de l'Information

Directeur de thèse : Basel Solaiman

Soutenue le 18 Décembre 2014

Jury :

M. Ali Khenchaf, Professeur, ENSTA Bretagne (Président)
M. Jean-Paul Haton, Professeur, Université Henri Poincaré, Nancy 1 (Rapporteur)
M. Kamel Hamrouni, Professeur, Ecole Nationale d'Ingénieurs de Tunis (Rapporteur)
M. Basel Solaiman, Professeur, Télécom Bretagne (Directeur de thèse)
M. Didier Guériot, Maître de Conférences, Télécom Bretagne (Examineur)
M. Zahzah El-hadi, Maître de Conférences, Université La Rochelle (Examineur)

Résumé

Cette thèse s'inscrit dans le contexte de la représentation et du traitement possibiliste des connaissances en vue de concevoir un système d'interprétation de scène observée par de multiples capteurs. Les capteurs fournissant une ou des informations imparfaites, les recherches se sont focalisées sur une représentation possibiliste des connaissances. En effet, ce formalisme permet d'introduire et de manipuler explicitement au sein d'un même cadre, certaines imperfections portées par les informations collectées mais aussi certaines connaissances *a priori* sur la scène, les capteurs, les conditions d'acquisition exprimées de manière vague soit par des experts soit résultant d'observations ou de jugements ambigus.

Nous avons donc adopté une approche basée sur la théorie des possibilités afin d'aborder ce sujet et ceci, dans un environnement multi-sources de connaissances : des capteurs, des sources additionnelles de connaissances, etc. Ce travail est divisé en plusieurs étapes : état de l'art et maîtrise de la théorie des possibilités, analyse hiérarchique/sémantique de scène, étude de l'exploitation de la théorie des possibilités dans différents niveaux sémantiques (sous-pixelique, pixelique, objet et scène) et finalement, application de l'approche proposée à des fins d'interprétation de scène.

L'objectif final de cette thèse est de développer un système d'interprétation de scène basé sur une démarche possibiliste. Cette approche permet d'effectuer une analyse hiérarchique de scène en se reposant sur deux processus : ascendant et descendant. Le processus ascendant permet d'accumuler l'évidence sur l'existence des régions (ou objets), tandis que le processus descendant permet de mettre en cause, par l'expert, le contenu informationnel des régions et objets identifiés dans le processus ascendant. En effet, ce système met en place les outils possibilistes nécessaires à l'exploitation conjointe de plusieurs sources de connaissances (fusion) afin d'améliorer l'interprétation des données acquises en vue d'une représentation plus riche de la scène observée : les deux principales applications visées sont donc la segmentation / classification (processus ascendant) et le démixage (processus descendant).

Les performances du système d'interprétation proposé sont évaluées en utilisant plusieurs types d'images (images de synthèse, images mammographiques, images du type IRM, images SONAR, et Images satellitaires). Les résultats obtenus sont très encourageants et montrent l'efficacité de la théorie des possibilités comme un cadre de représentation des connaissances *a priori* et comme outil de raisonnement permettant d'extraire de nouvelles connaissances à partir des connaissances disponibles. De plus, l'approche proposée s'est montrée très efficace pour l'intégration de plusieurs sources des connaissances, pour l'utilisation de la notion de la similarité afin d'avoir une idée plus approfondie du contenu thématique des régions (ou objets) identifiées dans la scène.

Mots clés :

Interprétation de scène, Théorie des possibilités, Raffinement de distribution de possibilités, Mesure de similarité, fusion de connaissances.

Abstract

This thesis considers the context of the possibilistic knowledge representation and processing in order to develop a scene interpretation system observed by multiple sensors. The research is focusing on a possibilistic knowledge modeling of the imperfect information provided by the sensors. Indeed, this formalism allows introducing and explicitly manipulating, within the same framework, certain imperfections carried by the collected information and also some *priori* knowledge about the scene, sensors, and acquisition conditions expressed by experts or resulting from observations or ambiguous judgments.

We, therefore, have adopted an approach based on the use of possibility theory to address this issue in a multi-source knowledge environment: sensors, additional sources of knowledge, etc. This work is divided into several stages: study of the state of the art and mastery of the possibility theory, hierarchical/semantic scene analysis, study of the application of possibility theory in different semantic levels (sub-pixel, pixel, region, object and scene), and finally the application of the proposed approach for scene interpretation.

The final objective of this thesis is to develop a scene interpretation system based on a possibilistic approach. This approach performs a hierarchical scene analysis relying on two processes: ascending and descending processes. The ascending process allows accumulating evidences on the existence regions (or objects), while the descending process allows calling into question, by the expert, the informational content of the regions and objects identified in the ascending process. Indeed, this system sets up the possibilistic tools allowing to jointly exploiting multiple sources of knowledge (fusion) in order to improve the interpretation of the acquired data and, thus, enriching the representation of the observed scene. The two main performed applications are segmentation/classification (ascending process) and the unmixing (descending process).

The performance of the proposed system is evaluated using several types of images (synthetic images, mammographic images, MRI images, SONAR images and satellite images). The obtained results are very encouraging and show the effectiveness of possibility theory as a framework for *priori* knowledge representation and as a reasoning tool to extract new knowledge from available knowledge. Moreover, the proposed approach has been very effective for integrating multiple knowledge sources, for using of the concept of similarity in order to get a deeper idea of thematic content of the regions (or objects) identified in the observed sense.

Keywords:

Scene interpretation, Possibility theory, Refining of possibility distribution, Similarity measure, Knowledge fusion.

Remerciements

Au terme de cette expérience riche et passionnante, je tiens à exprimer ma haute reconnaissance à mon directeur de thèse **Prof. Basel SOLAIMAN**, chef du département ITI (Images et Traitement de l'Information) à Télécom Bretagne, pour son encadrement, ses conseils, sa disponibilité, son soutien tout au long de l'élaboration de cette thèse.

Je remercie vivement **M. Didier GUEROT** pour son encadrement, sa générosité, sa patience et ses conseils durant la période dans laquelle il a supervisé mon travail.

Je voudrais exprimer toute ma reconnaissance à **l'Institut Supérieur de Sciences Appliquées et de la Technologie (ISSAT) en Syrie**, qui m'a financé tout au long de ce travail de recherche.

Mes remerciements s'adressent également à **M. Ali KHENCHAF**, Professeur à l'ENSTA Bretagne, pour l'honneur qu'il m'a accordé en présidant mon jury de thèse.

Je remercie les professeurs **M. Jean-Paul HATON** et **M. Kamel HAMROUNI** pour avoir accepté d'être les rapporteurs de ce travail.

Je remercie également **M. Zahzah EL-HADI** pour avoir accepté de participer à mon jury de thèse.

Enfin, un grand merci à toute l'équipe du département ITI, qui m'a accompagné au cours des années de thèse, et particulièrement **M. Shaban ALMOUAHED** et **Prof. Amer BAGHDADI**.

*Je dédie mes travaux de thèse
à ma chère femme Dima,
à mon cœur Christina,
et à toute ma famille.*

Sommaire

INTRODUCTION GENERALE.....	1
1. INTERPRETATION DE SCENE	7
1.1 Introduction.....	7
1.2 Cadre général d'interprétation de scène.....	8
1.3 Processus d'interprétation de scène	9
1.4 Méthodes de reconnaissance de formes	13
1.4.1 Les méthodes bayésiennes.....	13
1.4.2 Les méthodes markoviennes.....	13
1.4.3 Les modèles de mélanges	14
1.4.4 Les méthodes basées sur la théorie des ensembles flous	14
1.4.4.1 L'algorithme Fuzzy C-Means (FCM) :	15
1.4.4.2 Les systèmes à base de règles floues:.....	15
1.4.5 Les méthodes de fusion d'information multi-sources.....	16
1.5 Niveaux d'interprétation d'une scène	18
1.5.1 Le niveau sous-pixelique	19
1.5.2 Le niveau pixelique	20
1.5.3 Le niveau région.....	23
1.5.4 Le niveau objet	25
1.5.5 Le niveau scène	27
1.6 Conclusions.....	29
2. IMPERFECTION DE L'INFORMATION - CADRES MATHEMATIQUES	33
2.1 Introduction.....	33
2.2 Caractérisation d'information (imperfection et hétérogénéité)	33
2.2.1 Définition d'une information.....	34
2.2.2 Caractérisation d'une information	35
2.2.3 Notion d'information imparfaite.....	36
2.2.3.1 L'incertitude.....	36
2.2.3.2 L'ambiguïté.....	37
2.3 Approches de modélisation d'informations imparfaites	38
2.3.1 Théorie des probabilités.....	38
2.3.2 Théorie des fonctions de croyance	39
2.3.3 Théorie des ensembles flous	40
2.3.3.1 Définitions.....	40
2.3.3.2 Opérations sur les ensembles flous	41
2.3.4 Théorie des possibilités	43
2.3.4.1 Distribution de possibilités.....	43
2.3.4.2 Mesure de possibilité et mesure de nécessité	44
2.3.5 Règles de décision possibiliste	45
2.3.5.1 Règle de décision basée sur la possibilité maximale.....	45
2.3.5.2 Règle de décision basée sur la mesure de nécessité maximale.....	46
2.3.5.3 Règle de décision basée sur la maximisation de l'indice de confiance	46
2.4 Fusion d'information dans un cadre possibiliste.....	47
2.4.1 Opérateurs de fusion conjonctive	47
2.4.2 Opérateurs de fusion disjonctive	48

2.4.3	Opérateurs de fusion adaptative	49
2.5	Théorie des possibilités en interprétation de scène	49
2.6	Conclusion	50
3.	ESTIMATION DES DISTRIBUTIONS DE POSSIBILITES DANS LES SYSTEMES DE CLASSIFICATION PIXELIQUE.....	53
3.1	Introduction.....	53
3.2	Etat de l'art des méthodes d'estimation des distributions de possibilité	55
3.2.1	Méthodes d'estimation guidées par l'expert.....	55
3.2.1.1	Méthodes heuristiques.....	56
3.2.1.2	Méthodes hybrides	56
3.2.2	Méthodes de transformation d'une distribution de probabilités en une distribution de possibilités	57
3.2.2.1	Principe de cohérence probabilités-possibilités de Zadeh.....	58
3.2.2.2	Transformation de Dubois et Prade.....	58
3.2.3	Méthodes à base d'apprentissage.....	59
3.2.3.1	L'algorithme Fuzzy C-Means (FCM)	59
3.2.3.2	Méthode basée sur l'algorithme des k plus proches voisins.....	60
3.3	Approche proposée pour le raffinement des distributions de possibilités	61
3.3.1	Estimation des distributions de possibilités	62
3.3.2	Projection des connaissances possibilistes	63
3.3.3	Sélection des germes possibilistes	64
3.4	Expérimentations et résultats	67
3.4.1	Présentation des données expérimentales	67
3.4.2	Evaluation de la méthode de sélection des germes possibilistes	68
3.4.3	Effet de la largeur L du noyau de la méthode KDE.....	70
3.4.4	Comportement de l'approche proposée	73
3.4.5	Résultats expérimentaux sur des images médicales.....	75
3.5	Conclusions.....	76
4.	SIMILARITE POSSIBILISTE	79
4.1	Introduction.....	79
4.2	Mesures de similarité	80
4.2.1	Mesures de similarité métriques	81
4.2.2	Mesures de similarité ensemblistes	82
4.2.3	Mesures de similarité floues	83
4.3	Similarité et interprétation de scène.....	84
4.4	Similarité inter-classes	85
4.4.1	Exemple d'évaluation de la similarité entre deux classes.....	87
4.5	Similarité possibiliste et classification pixelique	89
4.5.1	Démarche de classification pixelique	89
4.5.2	Résultats de classification pixelique.....	90
4.5.2.1	Images simulées	91
4.5.2.2	Images réelles.....	93
4.6	Similarité possibiliste et démixage pixelique.....	94
4.6.1	Démixage possibiliste d'images simulées	97
4.6.2	Démixage possibiliste d'images réelles.....	100
4.7	Conclusions.....	101
5.	SYSTEME DE SEGMENTATION/INTERPRETATION POSSIBILISTE D'IMAGE.....	103
5.1	Introduction.....	103
5.2	Analyse ascendante pour la segmentation possibiliste.....	106
5.2.1	Raffinement-Estimation des distributions de possibilités.....	107
5.2.2	Diffusion des connaissances possibilistes	107
5.3	Comportement de la démarche de segmentation proposée	111
5.3.1	Evaluation des performances dans le cas de classes monomodales.....	111
5.3.2	Evaluation des performances au cas des classes multimodales	116
5.3.3	Evaluation des performances au cas d'images réelles	120
5.4	Proposition d'une démarche « descendante » de démixage.....	127
5.4.1	Application de l'approche descendante aux images réelles.....	132
5.5	Conclusion	134

6. VERS UN SYSTEME GLOBAL D'INTERPRETATION DE SCENE.....	137
6.1 Introduction.....	137
6.2 Fusion inter-primitives.....	140
6.3 Application de la phase ascendante à une image de télédétection	142
6.3.1 Le raffinement-estimation des distributions de possibilités.....	143
6.3.2 La projection des connaissances possibilistes	144
6.3.3 La diffusion des connaissances possibilistes et segmentation possibiliste	146
6.4 Phase descendante et analyse de scène en télédétection	147
6.4.1 Modélisation des régions	148
6.4.2 Similarité de régions	151
6.4.3 Rapport d'interprétation de scène.....	156
6.5 Conclusion	160
CONCLUSIONS ET PERSPECTIVES	161

Table des figures

Figure 0.1: Brique conceptuelle générique exploitant la richesse des connaissances portant sur des niveaux sémantiques voisins N-1, N et N+1 (avec N correspondant ici au niveau pixelique).....	3
Figure 1.1 : Le fossé sensoriel et le fossé sémantique.....	9
Figure 1.2 : Image d'une tumeur du sein par IRM.....	10
Figure 1.3 : Schéma général du processus d'interprétation de scène.....	11
Figure 1.4 : Architecture générale d'un système à base de règles floues.....	16
Figure 1.5 : Les différents niveaux de fusion des données multi-capteurs.....	18
Figure 1.6 : Méthode de segmentation d'image par diffusion des connaissances possibilistes.....	25
Figure 2.1 : Définition et structure de l'information.....	34
Figure 2.2 : Exemple d'une information en télédétection.....	35
Figure 2.3 : Formes classiques de fonctions d'appartenance.....	41
Figure 2.4 : Exemple de fonction d'appartenance trapézoïdale pour modéliser le concept "moyen".....	41
Figure 2.5 : Fonction d'appartenance de l'intersection de deux ensembles flous A et B.....	42
Figure 2.6 : Fonction d'appartenance de l'union de deux ensembles flous A et B.....	42
Figure 2.7 : Distribution de possibilités non-normalisée et degré d'inconsistance.....	44
Figure 2.8 : Indices de confiance associés aux différentes décisions (A_{n0} : décision ayant la possibilité maximale, A_{n1} : décision ayant la deuxième plus grande possibilité).....	47
Figure 2.9 : Opérateur de fusion conjonctive : \min	48
Figure 2.10 : Opérateur de fusion disjonctive : \max	49
Figure 2.11 : Opérateur de fusion adaptative de Dubois et Prade.....	49
Figure 3.1 : Méthode à base d'estimation statistique des paramètres.....	56
Figure 3.2 : Approche proposée pour le raffinement des distributions de possibilités.....	62
Figure 3.3 : Méthode d'estimation des distribution de possibilités à base de transformation probabilité-possibilité..	63
Figure 3.4 : Processus de classification pixelique des images.....	64
Figure 3.5 : (a) Images de synthèse avec les zones d'apprentissage (b) image de synthèse (cas de faible complexité de discrimination entre classes) et les distributions de possibilités initiales, (c) image de synthèse (cas de grande complexité de discrimination entre classes.) et les distributions de possibilités initiales.....	68
Figure 3.6 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil dans le cas gaussien: (a) cas de faible complexité de discrimination entre les classes (b) cas de grande complexité de discrimination entre les classes.....	69
Figure 3.7 : Effet de la valeur de seuil S sur les distributions de possibilités raffinées dans le cas de grande complexité de discrimination entre classes et pour S=0.1, 0.5 et 0.8.....	69
Figure 3.8 : (a)Image de synthèse (cas de grande complexité de discrimination entre classes.), (b) germes possibilistes sélectionnés pour la classe C_1 , (c) germes possibilistes sélectionnés pour la classe C_2	70
Figure 3.9 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de Rayleigh avec une faible complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale.....	71
Figure 3.10 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de Rayleigh avec une grande complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale.....	71
Figure 3.11 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de multimodales gaussien avec une faible complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale.....	72
Figure 3.12 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de multimodales gaussien avec une grande complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale.....	72
Figure 3.13 : (a) Image de synthèse (cas de faible complexité de discrimination), (b) les distributions de possibilités référence, initiale (c) les distributions de possibilités référence et raffinée, (d) Image de synthèse (cas de grande complexité de discrimination), (e) les distributions de possibilités référence, initiale (f) les distributions de possibilités référence et raffinée.....	74

Figure 3.14 : (a) Nombre de germes sélectionnés en cas de faible complexité de discrimination entre les classes et (b) nombre de germes sélectionnés en cas de grande complexité de discrimination entre les classes.....	74
Figure 3.15 : (Test-1) Image d'une tumeur du sein obtenue par imagerie IRM et (Test-2) une image mammographique	75
Figure 3.16 : Effet de raffinement des distributions de possibilités (a) et (d) image IRM et image mammographique respectivement avec deux zones d'apprentissage (b) et (e) distributions de possibilité initiales (c) et (f) distributions de possibilités raffinées.....	76
Figure 4.1 : Distance euclidienne D_E correspondant aux différentes mesures de similarité possibilistes, (a) et (c) deux distributions de possibilités gaussiennes représentant les deux classes C_1 et C_2 , (b) évolution de la mesure D_E en fonction de la différence des moyennes entre les classes C_1 et C_2 , (d) évolution de la mesure D_E en fonction de la différence d'écart-types entre les classes C_1 et C_2	88
Figure 4.2 : Distance euclidienne D_E correspondant aux différentes mesures de similarité possibilistes, (a) deux distributions de possibilités de Rayleigh représentant les deux classes C_1 et C_2 , (b) évolution de la mesure D_E en fonction de la différence de paramètre α_1 et α_2 entre les classes C_1 et C_2	88
Figure 4.3 : Similarité possibiliste et classification pixelique.	90
Figure 4.4 : (a) Image de synthèse originale, zones d'apprentissage et une zone autour du pixel considéré P_0 ; (b) distributions de possibilités π_1 , π_2 , et π_{P_0}	90
Figure 4.5 : Evolution du taux d'erreur de classification en fonction des caractéristiques des deux classes (moyennes et écarts-types) dans le cas Markovien (traits en pointillés) et dans le cas de l'approche proposée (traits pleins) pour des fenêtres de taille 3x3.....	91
Figure 4.6 : Evolution du taux d'erreur de classification en fonction des caractéristiques des deux classes (moyennes et écarts-types) dans le cas de l'approche proposée pour des fenêtres de taille 3x3 (a), 5x5 (b) et 7x7 (c)	92
Figure 4.7 : Evolution du taux d'erreur de classification en fonction des caractéristiques des deux classes (α_1 et α_2) (a) dans le cas Markovien et dans le cas de l'approche proposée pour des fenêtres de taille 3x3 et (b) dans le cas de l'approche proposée pour des fenêtres de taille 3x3, 5x5 et 7x7.....	92
Figure 4.8 : Classification d'image de synthèse (a) image originale, cas gaussienne et (d) image originale, cas Rayleigh ; (b) et (e) résultats de la classification par l'approche Markovienne (cas gaussien et cas Rayleigh respectivement) et ; (c) et (f) résultats de la classification par l'approche proposée (cas gaussienne et cas Rayleigh respectivement).....	93
Figure 4.9 : Classification d'image de mélanone (a) image originale et zones d'apprentissage, (b) estimation des distributions de possibilités des deux classes considérées, (c) classification par la méthode markovienne; (d) classification par l'approche proposée	94
Figure 4.10 : Schéma général d'un processus de démixage spectral utilisant le modèle linéaire.....	96
Figure 4.11 : Image de synthèse composée de deux classes C_1 et C_2 , zones d'apprentissage et mélange linéaire de C_1 et C_2	98
Figure 4.12 : (a) Image de synthèse composée de deux classes générées par deux distributions gaussiennes; (b) carte d'abondance de la classe C_1 et (c) carte d'abondance de la classe C_2	98
Figure 4.13 : (a) Image de synthèse composée de deux classes générées par deux distributions Rayleigh ; (b) carte d'abondance de C_1 et (c) carte d'abondance de C_2	99
Figure 4.14 : (a) Image mamographique de la base MIAS, (b) distributions de possibilités de tissus normaux et tumoraux estimées à partir des zones caractéristiques fixées par l'expert.....	100
Figure 4.15 : (a) Image mamographique segmentée par l'approche proposée, (b) les valeurs d'abondance de la classe « tumeur » et (c) la superposition des valeurs d'abondance de la classe « tumeur » et l'image originale	101
Figure 5.1 : Schéma global de l'approche proposée de segmentation/interprétation d'image basée sur une démarche possibiliste opérant dans les deux directions : ascendante et descendante	105
Figure 5.2 : Méthode de diffusion des connaissances possibilistes classe/classe pour une classe thématique donnée.....	108
Figure 5.3 : Méthode de diffusion croisée des connaissances possibilistes.....	108
Figure 5.4 : Configuration des sous-fenêtres (ou cliques) du filtre de Nagao	109
Figure 5.5 : Méthode de diffusion croisée de diffusion des connaissances possibilistes basée sur les cliques de Nagao.....	110
Figure 5.6 : Schéma global de l'analyse ascendante pour la segmentation possibiliste	111
Figure 5.7 : Images de synthèse (cas gaussien) (a) et (b) faible et grande complexité de discrimination entre les classes respectivement et (cas de Rayleigh) (c) et (d) faible et grande complexité de discrimination entre les classes respectivement.....	112
Figure 5.8 : Cartes possibilistes des classes C_1 et C_2 et les résultats de la segmentation à l'issue de la démarche de segmentation proposée (a) et (b) dans le cas gaussien avec une faible et grande complexité de discrimination entre les classes respectivement et (c) et (d) dans le cas de Rayleigh avec une faible et grande complexité de discrimination entre les classes respectivement.....	113
Figure 5.9 : Erreur quadratique moyenne entre deux cartes possibilistes successives dans la diffusion des connaissances possibilistes, (cas gaussien) (a) et (b) faible et grande complexité de discrimination	

entre les classes respectivement et (cas de Rayleigh) (c) et (d) faible et grande complexité de discrimination entre les classes respectivement.....	114
Figure 5.10 : (a), (b) et (c) Image de synthèse multimodales pour le cas de faible complexité de discrimination entre classes, distributions de possibilités initiales et raffinées respectivement des classes C_1 et C_2 et (d), (e) et (f) grande complexité de discrimination entre classes, distributions de possibilités initiales et raffinées respectivement des classes C_1 et C_2	117
Figure 5.11 : Cartes possibilistes des classes C_1 et C_2 et les résultats de la segmentation à l'issue de la démarche de segmentation proposée (a) et (b) dans le cas multimodales gaussien avec une faible et grande complexité de discrimination entre les classes respectivement	118
Figure 5.12 : Erreur quadratique moyenne entre deux cartes possibilistes successives dans la diffusion des connaissances possibilistes pour les classes multimodales gaussiennes (a) faible complexité de discrimination entre les classes et (b) grande complexité de discrimination entre les classes	118
Figure 5.13 : A gauche : Régions de lésion mammaire et, à droite : la vérité terrain associée (les zones tumorales sont délimitées par des radiologues)	121
Figure 5.14 : Deux images (Test-1 et Test-2) du type IRM du sein issues de la base RIDER) avec les zones d'apprentissage délimitées par l'expert. (a) et (c) les distributions de possibilités initiales (b) et (d) et des distributions de possibilités après raffinement	121
Figure 5.15 : Contour identifié de la zone tumorale et réalité terrain : (a) Méthode de segmentation par croissance de régions, (b) Méthode de segmentation semi-supervisée FPM, (c) Méthode de segmentation markovienne et, (d) Méthode de segmentation possibiliste proposée.....	122
Figure 5.16 : Deux images de test mammographiques avec les zones d'apprentissage (a) et (c) les distributions de possibilités initiales (b) et (d) les distributions de possibilités après raffinement.....	123
Figure 5.17 : Contour identifié de la zone tumorale: (a) Méthode de segmentation par croissance de régions, (b) Méthode de segmentation semi-supervisée FPM, (c) Méthode de segmentation markovienne et, (d) Méthode de segmentation possibiliste proposée.....	124
Figure 5.18 : Effet du processus de raffinement des distributions de possibilités sur les résultats de la segmentation possibiliste (a) segmentation par diffusion des connaissances et sans raffinement des distributions de possibilités, (b) segmentation par diffusion des connaissances et avec raffinement des distributions de possibilités, (c) Méthode de segmentation markovienne.	125
Figure 5.19 : Images SONAR utilisées pour l'évaluation de la démarche ascendante de segmentation possibiliste ..	126
Figure 5.20 : Image SONAR Test-1 avec zones d'apprentissage, (a) distributions de possibilités initiales des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances ; (b) distributions de possibilités raffinées des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances	126
Figure 5.21 : Image SONAR Test-2 avec zones d'apprentissage, (a) distributions de possibilités initiales des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances ; (b) distributions de possibilités raffinées des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances	127
Figure 5.22 : Schéma général du système de segmentation-classification possibiliste sous l'hypothèse d'un monde fermé.....	128
Figure 5.23 : Exemple de vecteur descriptif d'une région d'intérêt	129
Figure 5.24 : Image de synthèse-type avec les zones d'apprentissage	129
Figure 5.25 : Images de synthèse segmentées par l'approche possibiliste proposée : (a) cas de faible complexité de discrimination, et (b) cas de grande complexité de discrimination.....	130
Figure 5.26 : (a) Image de synthèse (cas de faible complexité de discrimination), (b) et (c) les cartes des taux d'abondance pour les deux classes C_1 et C_2 respectivement, (d) (e) les vecteurs descriptifs des taux d'abondance des classes pures (région R_1 correspondant à la classe C_1 , le fond de l'image correspondant à la classe C_2) (f) et (g) les vecteurs descriptifs des taux d'abondance des classes C_1 et C_2 dans la région d'intérêt R_2 (resp. régions R_3 et R_4)	131
Figure 5.27 : (a) (a) Image de synthèse (cas de grande complexité de discrimination), (b) et (c) les cartes des taux d'abondance pour les deux classes C_1 et C_2 respectivement, (d) (e) les vecteurs descriptifs des taux d'abondance des classes pures (région R_1 correspondant à la classe C_1 , le fond de l'image correspondant à la classe C_2) (f) et (g) les vecteurs descriptifs des taux d'abondance des classes C_1 et C_2 dans la région d'intérêt R_2 (resp. régions R_3 et R_4)	132
Figure 5.28 : Approche descendante appliquée pour l'interprétation d'images mammographiques (a) Image originale (cas de faible complexité de discrimination) ; (b) Image segmentée par la démarche possibiliste proposée ; (c) & (d) les cartes des taux d'abondance des deux classes « Tissu tumoral » et « Tissu sein » respectivement ; (e) Histogramme d'abondance de la classe Tissu tumoral de la zone suspecte ; (f) Histogrammes de taux d'abondance des deux classes : Tissu normal & Tissue tumorale (dans la zone de tissu sein) ; (g) Histogramme de taux d'abondance de la classe Tissu tumoral dans une fenêtre comportant les deux types de tissu	133

Figure 5.29 : Approche descendante appliquée pour l'interprétation d'images mammographiques (a) Image originale (cas de grande complexité de discrimination) ; (b) Image segmentée par la démarche possibiliste proposée ; (c) & (d) les cartes des taux d'abondance des deux classes « Tissu tumoral » et « Tissu sein » respectivement ; (e) Histogramme de taux d'abondance de la classe Tissu tumoral de la zone suspecte ; (f) Histogramme de taux d'abondance de la classe Tissu tumoral dans une masse suspecte	134
Figure 6.1 : Schéma général du système proposé d'interprétation de scène.....	139
Figure 6.2 : Phase ascendante du système proposé d'interprétation de scène et les deux stratégies de raffinement-estimation des distributions de possibilités adoptées : le raffinement-estimation des distributions de possibilités d'une façon individuelle (gauche) et le raffinement-estimation conjoint des distributions de possibilités (droite).	142
Figure 6.3 : (a) Image SPOT composée de quatre classes, (b) première bande, (c) deuxième bande, et (d) troisième bande	143
Figure 6.4 : Image du capteur SPOT et zones d'apprentissage considérées.....	143
Figure 6.5 : Distributions de possibilités pour les quatre classes dans la première bande (a) distributions de possibilités initiales, (b) distributions de possibilités raffinées image/image et (c) distributions de possibilités raffinées conjointement.....	143
Figure 6.6 : Distributions de possibilités pour les quatre classes dans la deuxième bande (a) distributions de possibilités initiales, (b) distributions de possibilités raffinées image/image et (c) distributions de possibilités raffinées conjointement.....	144
Figure 6.7 : Distributions de possibilités pour les quatre classes dans la troisième bande (a) distributions de possibilités initiales, (b) distributions de possibilités raffinées image/image et (c) distributions de possibilités raffinées conjointement.....	144
Figure 6.8 : Cartes possibilistes pour les quatre classes après le processus de la fusion (sans raffinement des distributions de possibilités), (a) image originale, (b) carte possibiliste de la classe C_1 , (c) carte possibiliste de la classe C_2 , (d) carte possibiliste de la classe C_3 , et (e) carte possibiliste de la classe C_4	145
Figure 6.9 : Cartes possibilistes pour les quatre classes après le processus de la fusion (après raffinement-estimation image/image des distributions de possibilités), (a) image originale, (b) carte possibiliste de la classe C_1 , (c) carte possibiliste de la classe C_2 , (d) carte possibiliste de la classe C_3 , et (e) carte possibiliste de la classe C_4	145
Figure 6.10 : Cartes possibilistes pour les quatre classes après le processus de la fusion (après raffinement-estimation conjoint des distributions de possibilités), (a) image originale, (b) carte possibiliste de la classe C_1 , (c) carte possibiliste de la classe C_2 , (d) carte possibiliste de la classe C_3 , et (e) carte possibiliste de la classe C_4	145
Figure 6.11 : Segmentation/classification de l'image SPOT sans raffinement des distributions de possibilités (a) image originale, (b) contours des régions identifiées dans l'image segmentée, (c) image classifiée.....	146
Figure 6.12 : Segmentation/classification de l'image SPOT au cas du raffinement-estimation image/image (a) image originale, (b) contours des régions identifiées dans image segmentée, (c) image classifiée.....	146
Figure 6.13 : Segmentation/classification de l'image SPOT au cas du raffinement-estimation conjoint (a) image originale, (b) contours des régions identifiées dans image segmentée, (c) image classifiée.....	147
Figure 6.14 : Classification de l'image SPOT au cas du raffinement-estimation conjoint (a) image originale, (b) image classifiée pour $S_0 = 0.5$, (c) image classifiée pour $S_0 = 0.1$ et (d) image classifiée pour $S_0 = 0$	147
Figure 6.15 : Système de modélisation des régions.....	148
Figure 6.16 : Application de la contrainte liée à la taille des régions détectées (a) image segmentée par raffinement-estimation conjoint et avant l'application de la contrainte $T_m=50$ pixels (b) image segmentée par raffinement-estimation conjoint et après l'application de la contrainte $T_m=50$ pixels.....	149
Figure 6.17 : Cartes d'abondance pour les quatre classes thématiques (a) image originale (b) carte d'abondance de la classe C_1 (c) carte d'abondance de la classe C_2 (d) carte d'abondance de la classe C_3 et (e) carte d'abondance de la classe C_4	149
Figure 6.18 : Exemple d'une région homogène-pure sur l'image SPOT : (a) Image originale segmentée sous la contrainte $T_m=50$ pixels, (b) La région homogène- pure, et (c) Le vecteur descriptif associé.....	150
Figure 6.19 : Exemple d'une région homogène-mixte comportant un fort mélange des quatre classes thématiques : (a) Image originale segmentée sous la contrainte $T_m=50$ pixels, (b) La région homogène-mixte, et (c) Le vecteur descriptif associé.....	150
Figure 6.20 : Exemple d'une région homogène-mixte comportant un fort mélange des quatre classes thématiques : (a) Image originale segmentée sous la contrainte $T_m=50$ pixels, (b) La région homogène-mixte, et (c) Le vecteur descriptif associé.....	150
Figure 6.21 : Exemple de régions identifiées en appliquant la phase ascendante et leurs vecteurs descriptifs (a) image segmentée, (b) Régions R_1 , R_2 , R_3 , R_4 et R_5 et (c) vecteurs descriptifs	153
Figure 6.22 : Exemple de l'effet de la normalisation sur les histogrammes d'abondance liée aux classes dans la région R_2 (a) les histogrammes d'abondance non normalisés et (b) les histogrammes d'abondance normalisés.....	154
Figure 6.23 : Exemple de construction d'un rapport d'interprétation de scène.....	158
Figure 7.1 : Brique conceptuelle générique exploitant la richesse des connaissances portant sur des niveaux sémantiques voisins N-1, N et N+1 (avec N correspondant ici au niveau objet).....	165

INTRODUCTION GENERALE

Dans le domaine de la vision par ordinateur et de l'intelligence artificielle, la construction d'un système d'interprétation de scènes constitue actuellement un champ très important et très prometteur de recherche. L'objectif visé par un tel système est d'apporter aux experts des outils d'aide à la décision leur facilitant l'exploitation conjointe de plusieurs sources de connaissances se rapportant à la scène observée et de pouvoir améliorer ainsi leur compréhension de cette même scène. Dans ce contexte, l'utilisation d'un système d'imagerie permet la projection de la scène, sous la forme d'images. On parle dans ce cas précis de systèmes d'interprétation d'image.

A. Problématiques de l'interprétation de scène

Un processus d'interprétation de scène rassemble les approches et les outils permettant l'identification du contenu sémantique des données collectées, données qui peuvent être issues d'une ou de plusieurs sources de connaissances. Ce processus a clairement besoin de la capacité d'extraire de nouvelles connaissances en s'appuyant sur les données collectées, tout comme le fait l'être humain, afin d'assurer l'interprétation de la scène observée. A ce titre, le contenu sémantique des données collectées à partir de la scène est donc fortement lié aux systèmes d'acquisition de l'information mais aussi aux connaissances *a priori* de l'expert qui donnent une sémantique aux observations. Ainsi, l'identification du contenu sémantique d'une scène est réalisée par la coopération entre les données collectées et les différents types de connaissances disponibles (connaissances *a priori* de l'expert, connaissances acquises au cours du processus d'interprétation, connaissances sur les processus physiques d'acquisition de l'information, etc.).

Cette coopération peut être réalisée sur chacun des cinq niveaux d'abstraction communément définis : le niveau sous-pixelique, le niveau pixelique, le niveau région, le niveau objet et le niveau scène globale. Ces cinq niveaux d'interprétation, basés essentiellement sur les méthodes de reconnaissance de formes, sont appelés *niveaux sémantiques d'interprétation*.

La description par l'expert (en se basant sur ses connaissances *a priori*) des informations et données collectées constitue un problème récurrent lors de la conception d'un système d'interprétation de scène. Ce problème devient plus complexe dans le contexte où les données sont très difficiles à discriminer (données bruitées ou fortement texturées, par exemple) ou lorsque le niveau sémantique d'interprétation s'élève (par exemple, le niveau sémantique « scène »). En effet, le processus d'interprétation de scène doit alors prendre en compte et intégrer dans ses mécanismes une richesse de connaissances *a priori* de plus en plus importante, car dans la majorité des systèmes d'interprétation de scène, les connaissances *a priori* de l'expert sont considérées comme représentatives du contenu sémantique à identifier. Néanmoins, ces connaissances *a priori* fournies par l'expert ne représentent généralement qu'un sous-ensemble des connaissances caractérisant complètement les contenus sémantiques de la scène. Ainsi, la capacité d'extraire de nouvelles connaissances à partir des données collectées et d'en qualifier leur certitude (essentielle lors du processus d'interprétation), constitue l'un des problèmes majeurs rencontrés pour la conception des systèmes d'interprétation de scène.

En général, les données collectées ainsi que les connaissances extraites au cours de ce processus souffrent d'un ou de plusieurs types d'imperfections telle que l'incertitude et l'imprécision. De plus, les connaissances *a priori* de l'expert sont généralement exprimées d'une manière vague. Ceci rend le processus d'interprétation de scène encore plus complexe. Pour ces raisons, une représentation adéquate des connaissances mises en jeu et des imperfections devient indispensable. Une telle représentation doit aussi permettre, d'un côté d'intégrer et d'exploiter conjointement les connaissances disponibles issues des différents types de sources de connaissance dans le processus d'interprétation et d'un autre côté, de manipuler explicitement les différentes formes d'imperfections entachant les données collectées.

De plus, le processus d'interprétation de scène est en général conditionné par l'objectif visé par l'expert. En effet, cet objectif permet de préciser le niveau sémantique d'interprétation de scène et de guider ainsi le processus d'interprétation afin d'extraire essentiellement la partie des connaissances intéressantes qui sont associées au niveau sémantique visé par l'objectif. De l'autre côté, la majorité des systèmes d'interprétation de scène (concernant les trois niveaux sémantiques d'interprétation de scène : le niveau sous pixelique, le niveau pixelique et le niveau région) opèrent sur le niveau pixelique pour aboutir au niveau sous pixelique ou pour aboutir au niveau région, sans formaliser les liens entre ces trois niveaux. Pourtant, un tel lien ne peut qu'améliorer les résultats globaux de l'interprétation.

Cette thèse s'inscrit donc dans le domaine des systèmes d'interprétation de scènes observées par de multiples capteurs, dans l'objectif de mettre à la disposition de l'expert un ensemble d'outils destiné à l'aider au cours de son processus de compréhension de la scène observée et de sa prise de décision.

Dans un premier lieu, ces outils permettent à l'expert, d'une manière simple et flexible de représenter ses connaissances liées aux informations et données collectées à partir de la scène et ceci, dans le contexte où ces informations et données sont très difficiles à discriminer tout en prenant en compte les différentes formes d'imperfections portées par ces données, informations et connaissances disponibles.

Dans un deuxième lieu, les outils informatiques offrent d'une part, la possibilité d'exploiter conjointement plusieurs sources des connaissances (opération référencée par le terme : fusion) en permettant notamment la production de nouvelles connaissances à partir des données disponibles et d'autre part, de formaliser des liens entre les niveaux d'interprétation considérés et ainsi, la remise en cause ou plus généralement le raffinement du contenu sémantique identifié.

B. Approche générique proposée

Dans ce travail de recherche, nous proposons un système de segmentation/interprétation d'image basé sur la théorie des possibilités dans un environnement multi-sources de connaissances pour les trois niveaux sémantiques d'interprétation suivants : sous-pixelique, pixelique, région. En effet, cette théorie offre plusieurs avantages au sein d'un processus d'interprétation de scène parmi lesquels l'efficacité et la simplicité en termes de représentation et d'intégration des connaissances *a priori* de l'expert, à des fins de traitement et de fusion au sein d'un même cadre, d'informations imparfaites issues de différentes sources de connaissances. Par ailleurs, ce système d'interprétation de scène repose sur une analyse hiérarchique qui opère sur trois niveaux d'interprétation voisins (ici pixelique, région et sous-pixelique) au travers de deux phases de traitement : une phase ascendante et une phase descendante (Figure 0.1).

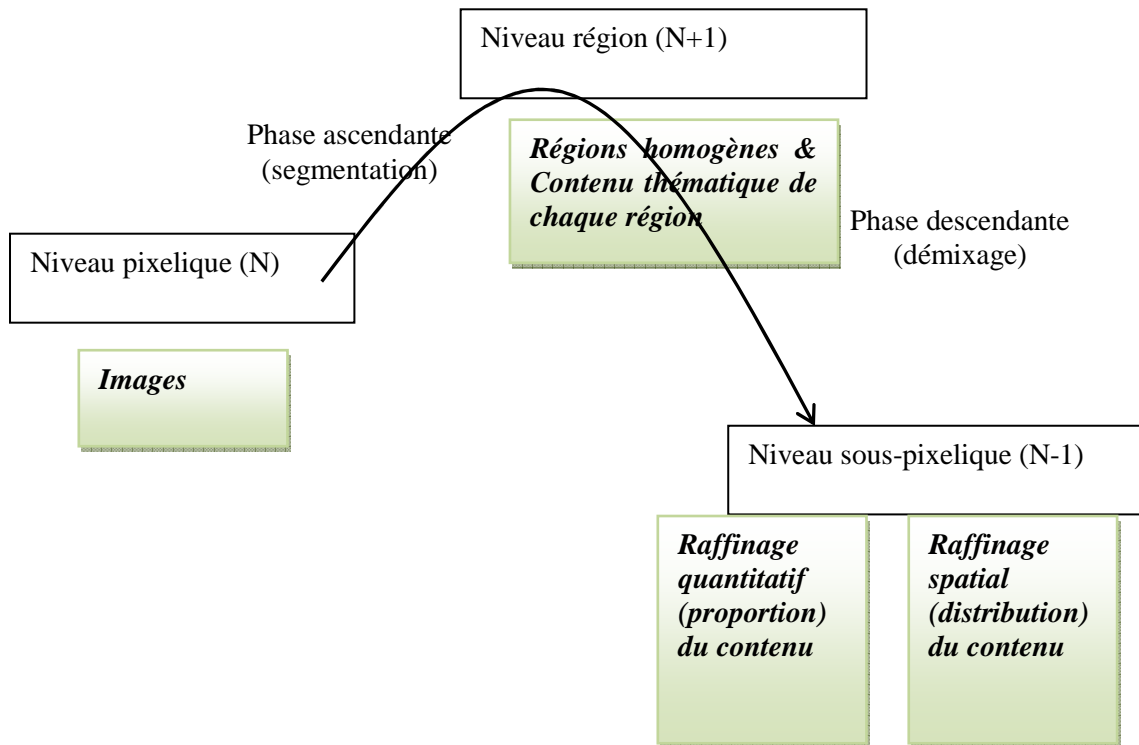


Figure 0.1 : Brique conceptuelle générique exploitant la richesse des connaissances portant sur des niveaux sémantiques voisins N-1, N et N+1 (avec N correspondant ici au niveau pixelique)

La phase ascendante a pour objectif de réaliser l'analyse et la segmentation d'image représentant une scène observée. En effet, cette phase, initialisée par les connaissances *a priori* de l'expert liées au contenu sémantique, comporte deux étapes consécutives : *le raffinement-estimation des distributions de possibilités* et *la diffusion des connaissances possibilistes*. L'étape du raffinement-estimation des distributions de possibilités permet la production de nouvelles connaissances à partir des connaissances disponibles en vue d'une représentation plus riche du contenu sémantique à identifier. En se basant sur l'ensemble des connaissances issues de la première étape, une étape de diffusion des connaissances possibilistes est menée afin d'accumuler les évidences sur les régions et ainsi, de finaliser le processus de segmentation de l'image. Les résultats ainsi obtenus se trouvent donc clairement au niveau sémantique « région » de par le fait des connaissances initiales apportées par l'expert et décrivant des régions de la scène.

La phase descendante a pour objectif d'améliorer l'interprétation des différentes régions homogènes identifiées dans la phase ascendante et donc de permettre de raffiner le contenu thématique de ces régions. Nous proposons dans cette phase, une approche de démixage pixelique basée sur la notion de la similarité possibiliste. Cette approche permet d'analyser l'ensemble des pixels composant chaque région identifiée afin de fournir une représentation plus riche du contenu thématique de ces régions (raffinage quantitatif permettant d'avoir le contenu thématique global dans la région considérée et/ou raffinement spatial permettant d'avoir la distribution spatiale de ce contenu thématique). Cette représentation permet aussi de réaliser, dans une étape ultérieure, une phase d'analyse de la similarité entre les différentes régions en fonction de leurs contenus thématiques dans l'objectif de fusionner les régions ayant des caractéristiques similaires et donc compléter l'analyse de la scène observée.

C. Plan du manuscrit

Ce manuscrit est donc organisé en six chapitres :

- Le premier chapitre vise à introduire le processus d'interprétation de scène tout en mettant l'accent sur les éléments essentiels et les problématiques concernées par notre travail de recherche (la conception d'un système d'interprétation de scène dans un contexte de multi-sources de connaissances). Un bref rappel de certaines méthodes de reconnaissance de formes utilisées dans les systèmes d'interprétation de scène est présenté. Une étude bibliographique des systèmes d'interprétation de scène appartenant à chacun des niveaux d'interprétation (niveau sous-pixelique, niveau pixelique, niveau région, niveau objet et niveau scène globale) est aussi donnée dans ce chapitre.
- Le deuxième chapitre est dédié à la définition de la notion d'information imparfaite ainsi qu'aux différents cadres mathématiques utilisés pour la représentation et le traitement de ces informations (théorie des probabilités, théorie des fonctions de croyance, théorie des ensembles flous et théorie des possibilités). Une attention particulière est portée à la théorie des possibilités et aux différents concepts de base de cette théorie : distribution de possibilités, mesure de possibilité, mesure de nécessité, opérateurs de fusion et critères possibilistes de décision. L'utilité et les raisons qui nous ont incitées au choix de la théorie des possibilités dans un processus d'interprétation de scène sont également détaillées.
- Le troisième chapitre aborde essentiellement la problématique de l'estimation des distributions de possibilités à partir d'un ensemble réduit de connaissances fournies par un expert. Dans un premier temps, un état de l'art sur les méthodes d'estimation des distributions de possibilités dans le domaine de l'interprétation de scène, est brièvement présenté. Dans un deuxième temps, et en se basant sur une estimation initiale des distributions de possibilités, nous proposons et évaluons une approche qui exploite le contexte spatial d'une façon itérative, afin d'enrichir les connaissances exprimées par l'expert et réaliser ainsi un raffinement de l'estimation des distributions courantes de possibilités.
- Le quatrième chapitre est consacré à la notion de similarité entre deux informations possibilistes (i.e. deux distributions de possibilités). Après avoir présenté les différentes catégories existantes de mesures de similarité, nous proposons une méthode permettant de choisir une mesure de similarité, en se basant sur la capacité de discrimination entre les classes thématiques contenues dans la scène observée. Ensuite, nous proposons une démarche méthodologique permettant de réaliser la classification pixelique mais aussi l'opération de démixage pixelique en se basant sur le concept de la similarité possibiliste spatiale. Cette démarche sera appliquée et évaluée dans le contexte d'un mono-capteur.
- Le cinquième chapitre est dédié à la mise en œuvre de l'approche possibiliste (en deux phases, ascendante et descendante) de segmentation/interprétation d'image proposée. La phase ascendante (qui réalise la segmentation d'image) regroupe deux étapes : le *raffinement* et la *diffusion des connaissances possibilistes*. Une attention particulière est portée à l'effet d'enrichissement des connaissances d'expert par de nouveaux échantillons lors du processus de segmentation dans un contexte où les classes sont difficiles à discriminer (faible contraste,

environnement très bruité ou l'existence de classes multimodales). A l'issue de la phase descendante, chaque région identifiée dans la phase ascendante sera caractérisée par un vecteur descriptif composé des histogrammes des taux d'abondance aux différentes classes thématiques présentes dans cette région (démixage).

- Le sixième chapitre propose une extension de l'approche de segmentation/interprétation d'image, proposée dans le cinquième chapitre. L'objectif est d'exploiter plusieurs primitives dans la phase ascendante, et de quantifier la similarité (en termes de contenu thématique) entre les régions identifiées dans la phase ascendante afin de réaliser la fusion entre ces régions. Une évaluation globale de l'approche proposée est aussi détaillée dans ce chapitre.

Finalement, nous concluons en présentant les objectifs atteints dans ce travail et en évoquant de nouvelles perspectives de futures recherches pour améliorer le système d'interprétation proposé.

Chapitre 1

INTERPRETATION DE SCENE

1.1 Introduction

Le terme *scène* désigne un sous-ensemble d'un environnement au sein duquel sont considérés de multiples objets ayant un sens sémantique et des relations spatiales reliant ces objets. Le processus d'interprétation de scène a pour objectif d'identifier et d'analyser ce contenu sémantique afin d'extraire de nouvelles connaissances.

Généralement, deux types de connaissance sont nécessaires pour mener à bien ce processus : les connaissances issues de la scène elle-même (par exemple, les informations collectées à partir d'un patient en analyse d'une scène médicale, des mesures de réflectance en analyse d'une scène en télédétection, ...) et des connaissances liées à l'expert impliqué dans ce processus d'interprétation (le médecin, dans le cas de l'analyse d'une scène médicale ou le photo-interprète opérant sur une scène d'environnement). La coopération entre ces deux formes de connaissances permet de révéler, de comprendre et d'interpréter le contenu de la scène.

L'utilisation d'un capteur d'imagerie, projette la scène réelle dans un nouvel espace de représentation des connaissances, celui d'une image numérique. En d'autres termes, grâce au système d'imagerie, l'espace réel 3D est transformé en un espace 2D comportant des mesures physiques qui constituent l'image. Pour cette raison, les deux termes d'interprétation de scène et d'interprétation d'image sont souvent confondus.

Notons que l'échelle d'analyse de la scène observée par le système d'imagerie peut s'étendre du niveau "microscopique" (un prélèvement de laboratoire observé par un microscope) jusqu'à un niveau "macroscopique" (une ville observée par un satellite). L'image représentant la scène est considérée comme une matrice 2D dont les éléments (pixels) sont des vecteurs de mesures sur un ensemble de primitives « *features* » sélectionnées. Ces primitives peuvent être par exemple, des valeurs numériques (DN : Digital Numbers) mesurant les informations photométriques d'une scène perçue dans l'une des

bandes spectrales utilisées par le système d'imagerie (« *spectral patterns* »), ou des caractérisations de relations spatiales (i.e. des mesures contextuelles et texturales opérant sur des « *spatial patterns* »), ou encore des primitives temporelles (« *temporal patterns* »).

Plusieurs images d'une même scène peuvent être capturées par plusieurs types de systèmes d'imagerie simultanément (seules les dimensions spatiales et spectrales sont considérées) mais aussi éventuellement à des moments différents. Aux deux dimensions précédentes s'ajoute donc, une dimension temporelle.

Le terme "interprétation de scène" désigne les mécanismes de raisonnement et d'analyse visant à extraire la sémantique contenue dans ces images, tout comme le font les humains [1]. Etant donné que ce processus opère sur une « projection » (et donc une restriction) de la scène réelle dans un espace d'images, au travers de capteurs physiques, cette tâche d'interprétation exige l'intégration de connaissances *a priori* liées à l'expert ainsi que de connaissances contextuelles diverses, afin d'analyser la scène.

Dans ce chapitre, nous abordons le cadre général de l'interprétation de scène. Ensuite, nous présentons le processus d'interprétation de scène. Dans la troisième partie, nous donnons un bref rappel de certaines méthodes de reconnaissance de formes souvent utilisées en interprétation de scène. Nous enchaînons par la présentation des niveaux sémantiques du processus d'interprétation en détaillant certains systèmes d'interprétation de scène appartenant à chacun de ces niveaux et ceci dans l'objectif de mettre en avant les problématiques liées à la conception d'un système d'interprétation de scène et qui nous intéressent dans notre travail.

1.2 Cadre général d'interprétation de scène

Dans le domaine de la vision par ordinateur et de l'intelligence artificielle, la question d'interprétation de scène est fortement liée à la perception visuelle, c'est-à-dire à la perception du monde réel par des systèmes d'imagerie (système visuel humain, caméra CCD, ...etc.). La notion de perception visuelle désigne l'ensemble des actes de détection d'une scène (les objets visibles, les structures et les événements), de reconnaissance et de description avec des symboles [1]. Bien que les humains arrivent facilement et de manière robuste à réaliser une telle perception visuelle, cette tâche reste toujours un défi pour les systèmes de vision par ordinateur ainsi que pour les approches d'intelligence artificielle.

La vision par ordinateur vise à imiter la vision humaine en utilisant ses propres mécanismes d'observation, d'analyse et d'interprétation de scène. Ainsi, son objectif est de doter les ordinateurs d'un système de vision permettant de comprendre leur environnement à partir des données issues de systèmes d'imagerie. Cette manière de procéder s'effectue par une série de procédures de traitement de données et d'extraction d'informations basées sur la géométrie, la physique, les statistiques et les théories d'apprentissage [2]. Chaque étape d'un tel système de traitement est conditionnée par des connaissances *a priori* spécifiques au domaine considéré [3].

Selon le paradigme de Marr [4], un système de vision est une succession de procédures qui permettent de transformer l'information d'un niveau d'abstraction vers un niveau d'abstraction plus élevé. Marr a ainsi organisé cette succession en trois niveaux de représentation : le bas niveau qui vise à détecter les changements d'intensité significatifs dans une image et à en extraire des caractéristiques de base (régions, contours,...etc.), le niveau intermédiaire qui décrit les liens entre les deux espaces 2D (image) et 3D (monde réel), et le haut niveau qui traite de la description complète d'une scène. Dans ce paradigme, Marr ne tient pas compte du rôle des connaissances *a priori* ou des connaissances d'expert pour la coopération entre ces trois niveaux [1]. Ainsi, l'extraction du contenu sémantique de la scène devient très difficile [6].

Aujourd'hui, et malgré le fait que les ordinateurs possèdent une puissance de calcul supérieure à l'être humain, on n'arrive pas à produire des systèmes de vision robustes. Cela est dû principalement aux deux fossés entre la réalité et ce qui est perçu (Figure 1.1) :

- Le *fossé sensoriel* qui est défini comme « le fossé dû à l'enregistrement, ou la projection, de la scène 3D sur une ou plusieurs images 2D » [5] [6].
- Le *fossé sémantique* qui est plus délicat à traiter. Il est défini comme « l'insuffisance des informations extraites à partir des pixels, pour interpréter la scène » [5] [6]. Il s'agit du manque de connaissances *a priori* nécessaires pour extraire la sémantique contenue dans l'image.

L'homme arrive pourtant, à combler ces deux fossés grâce à plusieurs avantages qu'il possède : beaucoup de connaissances *a priori*, sa faculté d'apprentissage, sa faculté d'approximation et d'analogie,°... A l'inverse, la machine, qui ne possède pas de tels avantages, a toujours des difficultés pour remédier à ce problème. Ceci explique le fait que l'humain et la machine peuvent avoir des interprétations différentes de la scène analysée.

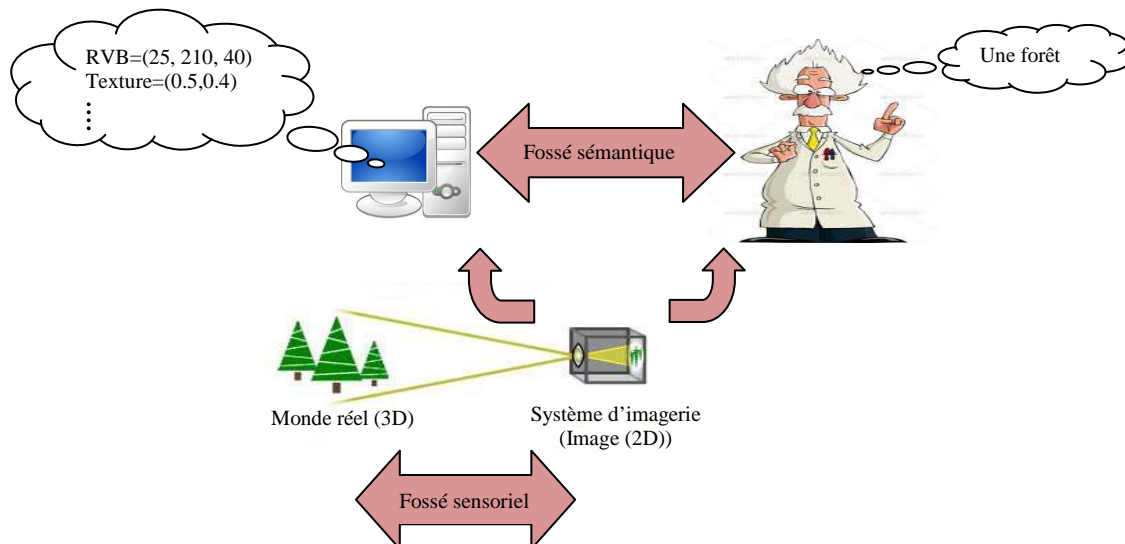


Figure 1.1 : Le fossé sensoriel et le fossé sémantique

Pour résumer, le processus d'interprétation de scène consiste à apporter des solutions aux problèmes liés à ces deux fossés. Dans la suite de ce chapitre, nous allons présenter ce processus tout en mettant en avant ses éléments pertinents.

1.3 Processus d'interprétation de scène

Le processus d'interprétation de scène vise à identifier le contenu sémantique de la scène observée [7] par l'exploitation conjointe des mesures numériques issues du système d'imagerie et de certaines connaissances *a priori* sur la scène analysée, sur les capteurs disponibles et les conditions d'acquisition exprimées de manière plus vague (soit par des experts ou résultant d'observations et de jugements). Cette coopération conjointe peut être réalisée aux différents niveaux d'interprétation : le niveau sous-pixelique, le niveau pixelique, le niveau région, le niveau objet ou le niveau scène globale. Ces niveaux d'interprétation, basés essentiellement sur les méthodes de reconnaissance de formes, sont appelés les niveaux sémantiques de l'interprétation de scène. La difficile tâche d'identification du contenu pour chaque niveau sémantique d'interprétation est fortement influencée par plusieurs facteurs :

- L'objectif final : les résultats de l'interprétation dépendent fortement de l'objectif recherché par l'expert. Cet objectif permet de préciser le niveau sémantique de l'interprétation et d'adopter une stratégie qui focalise l'analyse uniquement sur le contenu sémantique correspondant à ce niveau. Une telle stratégie permettrait d'exploiter une grande masse de données dans un temps raisonnable (par exemple, la fouille de données de télédétection) [15].
- Les connaissances *a priori* : qui peuvent être liées à l'expert ou extraites durant le processus d'interprétation (connaissances extraites à un instant donné). Ces connaissances jouent un rôle primordial dans le processus de l'interprétation. Chaque niveau sémantique du processus d'interprétation nécessite certains types de connaissances *a priori*. D'une manière générale, les connaissances *a priori* que l'on doit prendre en compte dans le processus d'interprétation d'une scène peuvent être classifiées en trois catégories [8] :
 - Les connaissances liées à la scène : modèles géométriques des objets dans la scène et leurs relations, le domaine d'application, etc.
 - Les connaissances portant sur la correspondance entre la scène et l'image : les propriétés des images, en particulier leur géométrie et le bruit dû au processus d'acquisition, sont liées au type du système d'imagerie (optique, radar, sonar, etc.) et à ses caractéristiques physiques (la longueur focale, l'angle d'acquisition des caméras, les résolutions spatiale et spectrale, la polarisation du capteur, etc.).
 - Les connaissances liées à l'image : types et relations spatiales entre les caractéristiques de base de l'image tels que les bords, les lignes, les régions, etc.

Les connaissances liées au domaine d'application ainsi qu'aux capteurs physiques d'acquisition des images peuvent influencer la manière dont la scène est perçue. A titre d'exemple, sans avoir de telles connaissances, l'image présentée d'une tumeur du sein dans la Figure 1.2 obtenue par imagerie IRM « *imagerie par résonance magnétique* » (Breast MRI dataset RIDER (<http://www.cancerimagingarchive.net/>)), peut également s'interpréter très différemment comme par exemple, la présence de plantes dans une image de télédétection.

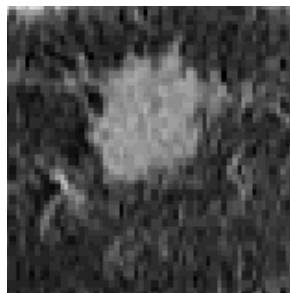


Figure 1.2 : Image d'une tumeur du sein par IRM

- La représentation des connaissances et des imperfections: cette représentation doit permettre, d'un côté d'intégrer les connaissances disponibles dans le processus d'interprétation et d'un autre côté, de gérer explicitement les différentes formes d'imperfections entachant les données aux différents niveaux sémantiques.
- Les méthodes qui réalisent l'interprétation: les méthodes utilisées aux différents niveaux d'un système d'interprétation de scène doivent être gérées d'une manière intelligente afin d'anticiper les effets d'imprécision [46]. Elles peuvent être sélectionnées *a priori* par l'expert (d'une façon

statique) et/ou durant le processus d'interprétation (d'une façon dynamique). Des telles méthodes accordent au processus d'interprétation de scène la capacité d'obtenir de bonnes performances dans des environnements très complexes [100]. Les connaissances *a priori* liées aux caractéristiques des données délivrées par le système d'imagerie telles que les résolutions spectrale, spatiale et temporelle, la polarisation et l'angle d'incidence jouent un rôle très important dans le traitement intelligent des images. Ces connaissances permettent la sélection convenable tant des données disponibles que des méthodes afin d'atteindre l'objectif visé [12] : par exemple, l'utilisation de données haute résolution telles que les images optiques SPOT-4 ou radar ERS-2 et des méthodes de fusion multi-sources de ces images, pour une analyse précise de la couverture terrestre [29].

Il est important de noter que le processus d'interprétation de scène peut être basé sur des données délivrées par un seul système d'imagerie ou sur des données délivrées par plusieurs systèmes d'imagerie, c'est-à-dire un système multi-capteurs permettant de fournir plusieurs projections physiques de la scène imagée. La Figure 1.3 illustre le schéma général du processus d'interprétation.

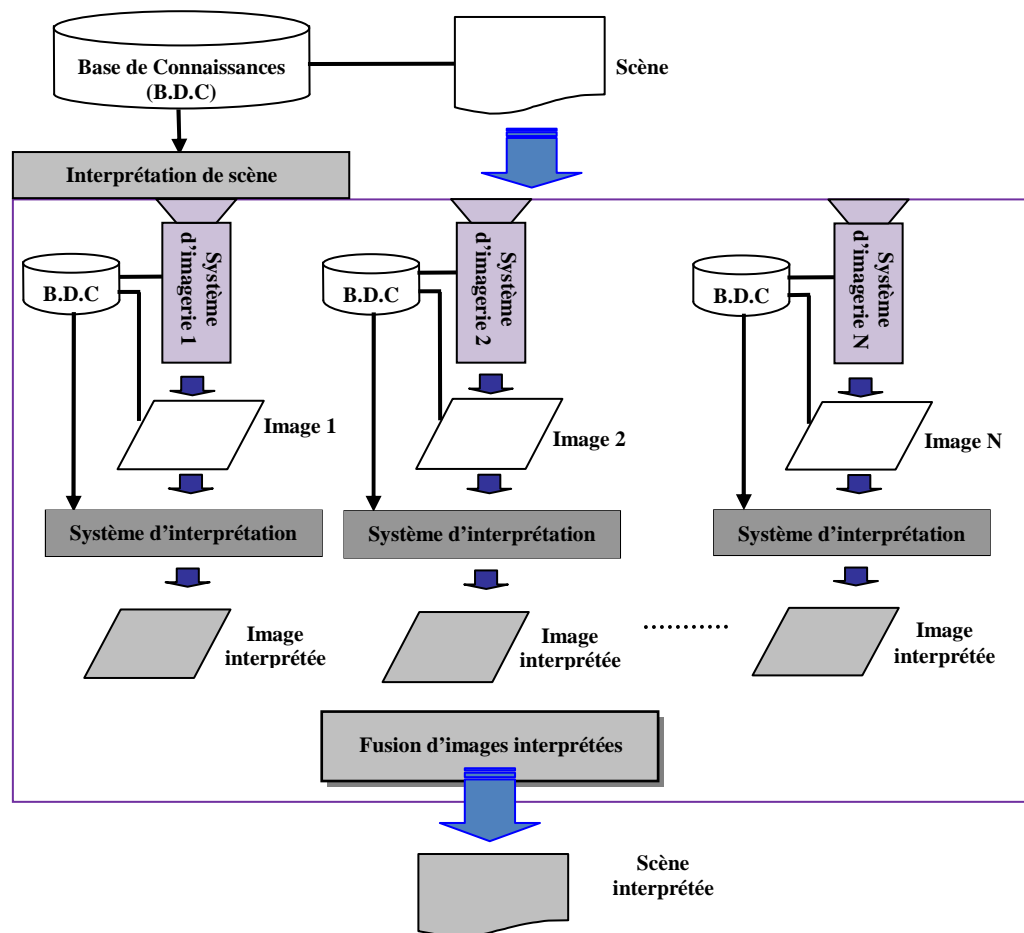


Figure 1.3 : Schéma général du processus d'interprétation de scène

Dans ce schéma, nous pouvons noter que l'interprétation de scène est souvent réalisée au niveau de chaque système d'imagerie, seul, le résultat final de l'interprétation étant obtenu par la fusion des différentes interprétations partielles. Cette approche présente l'avantage de pouvoir appliquer les connaissances *a priori* au niveau de chaque capteur d'une façon indépendante des autres capteurs. Ceci

est considéré comme une lacune aussi car l'information « croisée » porte une connaissance utile pour le processus d'interprétation.

Les systèmes de vision par ordinateur sont caractérisés par de nombreux types d'imperfections qui accompagnent leurs différents niveaux, depuis le traitement d'images (bas niveau) jusqu'aux analyses (niveau intermédiaire), la reconnaissance de formes et l'interprétation de scène (haut niveau) [46]. Par conséquent, un processus d'interprétation de scène doit posséder les mécanismes permettant de gérer explicitement les informations imparfaites pour que l'identification du contenu sémantique soit plus fiable. Afin de prendre en compte le problème des imperfections au cours d'un processus d'interprétation de scène, quatre cadres mathématiques sont couramment utilisés : la théorie des probabilités, la théorie des croyances, la théorie des ensembles flous et la théorie des possibilités [9]. Chacune de ces approches possède ses propres modes de représentation de connaissances et ses propres outils de traitement et de raisonnement appropriés pour les différents types d'imperfection d'information [11]. Généralement, la théorie des probabilités est la plus utilisée pour traiter la forme particulière de l'imperfection qui est l'incertitude. Cependant, il existe certains cas dans lesquels les sources d'imperfections ne sont pas simplement représentables par un modèle probabiliste tels que le manque de connaissances concernant la qualité de l'image (la définition de la qualité de l'image elle-même est basée sur la perception humaine, et diffère d'une personne à l'autre), l'imprécision dans les calculs, l'ambiguïté dans la description des objets, ... [46]. Dans ces cas, la théorie des croyances (aussi appelée théorie de Dempster-Shafer), la théorie des ensembles flous et la théorie des possibilités sont mieux adaptées pour traiter ces types d'imperfections.

La théorie des croyances représente l'appartenance de chaque pixel, région ou objet aux différentes classes dans la scène via des fonctions de masse de croyance. Ce mode de représentation permet d'intégrer, à la fois, l'imprécision et l'incertitude relatives aux connaissances disponibles.

La théorie des ensembles flous utilise des fonctions d'appartenance permettant la représentation de l'appartenance partielle et répondant ainsi à la non-exclusivité des connaissances. Ce mode de représentation floue permet donc, de modéliser l'imprécision dans un contexte ambigu.

Finalement, la théorie des possibilités permet la modélisation d'une forme particulière d'imperfection qui est celle d'une incertitude où les connaissances disponibles pour décrire cette incertitude sont données ou formulées d'une manière ambiguë (voir chapitre 2).

Notons que le cadre probabiliste est très limité en ce qui concerne la représentation et l'intégration des connaissances de l'expert, ce qui rend le processus d'interprétation de scène très difficile dans un tel cadre [46], alors que les autres cadres apportent des mécanismes permettant d'une façon flexible et satisfaisante de représenter et d'utiliser l'ensemble de ces connaissances.

Le processus d'interprétation de scène est basé essentiellement sur les méthodes de reconnaissance de formes. Plusieurs méthodes de reconnaissance de formes ont été développées afin d'améliorer les performances des systèmes de vision par ordinateur telles que les techniques basées sur l'intelligence artificielle, les systèmes experts [118], les réseaux de neurones [24] [25], les systèmes à base de règles [26], etc.

Dans la suite, nous allons donner un bref rappel des méthodes de reconnaissance de formes les plus utilisées dans les systèmes de vision par ordinateur. En particulier, nous nous intéressons aux méthodes utilisées dans le cadre des applications suivantes : le démixage, la classification et la segmentation.

Dans le cadre de notre travail, les applications que nous considérons se focalisent sur les trois premiers niveaux sémantiques d'interprétation de scène : le niveau sous pixelique, le niveau pixelique et le niveau région.

1.4 Méthodes de reconnaissance de formes

Jusqu'à la moitié des années 1990, les méthodes appliquées pour la reconnaissance de formes ont été principalement basées sur des approches statistiques classiques, telles que les méthodes bayésiennes [63] ou les méthodes basées sur le calcul de la distance minimale [64]. Bien que ces méthodes traditionnelles aient de bonnes performances, leur capacité générale à résoudre la confusion interclasse est limitée. Par conséquent, ces dernières années, et à la suite des progrès en technologie informatique, des stratégies alternatives ont été proposées, notamment l'utilisation des réseaux de neurones artificiels [65], les arbres de décision [66], la machine à vecteur de support [67], les méthodes issues de la théorie des ensembles flous [95]. L'incorporation des informations liées aux structures spatiales (telles que la texture, etc.) a aussi été proposée [20].

Dans la suite, on présente brièvement les méthodes de reconnaissance de formes les plus fréquemment utilisées pour les trois premiers niveaux sémantiques d'interprétation de scène : sous pixelique, pixelique et région.

1.4.1 Les méthodes bayésiennes

Le théorème de Bayes [63] a une influence longue et profonde sur la représentation statistique des connaissances. Deux éléments clés constituent la formule bayésienne : une distribution de probabilités *a priori* et une distribution de probabilités conditionnelles de vraisemblance. Par la combinaison de ces deux distributions, la probabilité conditionnelle *a posteriori* peut être exprimée de la façon suivante :

$$\Pr(C_m / x_n) = \frac{\Pr(x_n / C_m) \cdot \Pr(C_m)}{\Pr(x_n)} \quad (1.1)$$

où x_n est la valeur numérique mesurée de la primitive du « pattern » caractérisant le pixel, la région, ou même l'objet observé. Ce pattern est supposé appartenir à une classe thématique C_m issue d'un ensemble exhaustif et exclusif de M classes contenues dans la scène. $\Pr(C_m)$ est la distribution de probabilités *a priori* (la distribution de l'occurrence de la classe C_m) et $\Pr(x_n / C_m)$ est la distributions de probabilités conditionnelles aux différentes classes thématiques (distribution de vraisemblance). La règle de décision associe chaque pattern observé à la classe thématique pour laquelle la probabilité *a posteriori* $\Pr(C_m / x_n)$ est la plus élevée :

$$x_n \in C_{m_0} \text{ si et seulement si } \Pr(C_{m_0} / x_n) = \max\{\Pr(C_m / x_n), 1 \leq m \leq M\} \quad (1.2)$$

En supprimant le terme commun $\Pr(x_n)$, cette règle de décision, appelée maximum *a posteriori* (MAP), est réduite à la maximisation le produit des deux distributions $\Pr(C_m)$ et $\Pr(x_n / C_m)$.

1.4.2 Les méthodes markoviennes

Dans tout processus d'interprétation de scène, le contexte spatial apporte toute une gamme d'informations utiles. L'utilisation appropriée de la notion de contexte spatial permet l'élimination des ambiguïtés possibles, la récupération d'informations manquantes et la correction d'erreurs [18]. L'utilisation du contexte spatial pour modéliser les connaissances *a priori* (la distribution de probabilités *a priori* $\Pr(C_m)$) afin d'aider à l'interprétation de scène, est considérée comme une procédure raisonnable. En effet, un pixel est généralement entouré par des pixels de la même classe et il est donc moins probable d'avoir des pixels voisins appartenant aux autres classes de la scène.

Les champs de Markov constituent un outil utile et performant pour caractériser l'information relative au contexte spatial, outil qui a été largement utilisé pour la segmentation et la restauration d'images [19][20]. Ces champs sont des configurations (des cliques) définies sur le champ d'étiquettes C entourant chaque pixel. Le champ d'étiquettes C est obtenu à partir d'une étape d'initialisation par un expert ou par

une segmentation préliminaire de l'image considérée (par exemple, par seuillage d'histogramme ou l'application de l'algorithme des k-moyennes).

Une fois le contexte C modélisé d'une façon probabiliste (au moyen des champs de Markov), cette modélisation sera considérée comme une source de connaissances *a priori*. Le théorème de Bayes sera utilisé pour combiner ces connaissances *a priori* et l'ensemble des distributions de probabilités conditionnelles aux différentes classes thématiques. On retrouve alors une formule similaire à l'équation (1.1) :

$$\Pr(C/x_n) = \frac{\Pr(x_n/C) \cdot \Pr(C)}{\Pr(x_n)} \quad (1.3)$$

Ainsi, le problème se "résume" au calcul de la configuration optimale du champ d'étiquettes C qui maximise la probabilité *a posteriori* $\Pr(C/x_n)$. Pour cette optimisation, un algorithme itératif tel que l'ICM « *Iterated Conditional Modes* » [21], pourra être appliqué.

1.4.3 Les modèles de mélanges

Dans les modèles de mélanges, fréquemment utilisés en classification, on considère que les pixels formant l'image suivent une loi dont la fonction de densité est un mélange de plusieurs densités de probabilités (densité multimodale). Les modes dans un histogramme indiquent l'existence de plages de vecteurs d'attributs pour lesquels le nombre de pixels ayant ces vecteurs est plus élevé que sur d'autres plages. Ces modes peuvent correspondre à certaines régions ou caractéristiques (certains types de tissu, par exemple) dans l'image.

Soit x_n la valeur numérique mesurée de la primitive caractérisant un pixel dans une bande spectrale considérée (e.g. niveau de gris). Sa distribution $f(x_n)$ peut être représentée par une distribution de probabilités de la forme :

$$f(x_n) = \lambda_1 \cdot f_1(x_n) + \lambda_2 \cdot f_2(x_n) + \dots + \lambda_k \cdot f_k(x_n) \quad (1.4)$$

avec $0 \leq \lambda_i \leq 1$, $i = 1, 2, \dots, k$ et $\lambda_1 + \lambda_2 + \dots + \lambda_k = 1$. Les paramètres λ_i sont appelés *les coefficients de pondération* du mélange (*the mixing weights*). Le coefficient λ_i peut être interprété comme étant la probabilité *a priori* d'observer la valeur x_n sachant l'occurrence de la classe C_i ($\Pr(x_n/C_i)$).

Les fonctions f_i sont appelées les densités du mélange. Bien que n'importe quelle loi puisse être suivie par chaque densité f_i , la plus courante et la plus simple à appliquer (et la plus rapide à calculer) est la loi normale dont la fonction de densité est une gaussienne (*Gaussian Mixture Model*) [22]. Les paramètres de chaque gaussienne (sa moyenne et sa variance) sont estimés en utilisant des méthodes telles que la méthode de maximum de vraisemblance. L'optimisation de cette méthode est souvent réalisée via l'algorithme Espérance-Maximisation (EM) qui est une méthode itérative pour optimiser la fonction de vraisemblance dans le cas de manque d'information [23].

1.4.4 Les méthodes basées sur la théorie des ensembles flous

L'un des problèmes majeurs rencontrés en reconnaissance de formes est le problème de la confusion inter-classes [63]. Ce problème, qui constitue un environnement d'ambiguïté, peut se produire pour les scènes ayant des motifs complexes, c'est-à-dire avec un chevauchement spectral et/ou spatial des classes. Dans le domaine spectral, les caractéristiques de réflectance, émission, ou rétrodiffusion de différentes classes peuvent être similaires. Dans le domaine spatial, un pixel, une région, ou même un objet peut contenir des zones représentatives de plus d'une classe (i.e. problème de pixels mixtes dits « *mixels* »). En effet, dans les méthodes de reconnaissance de formes basées sur la représentation probabiliste, chaque pixel, région ou objet est supposé appartenir à une et une seule classe parmi les différentes classes

figurantes dans la scène. Par conséquent, ces méthodes ne fournissent pas les mécanismes permettant de faire face au problème de « mélange » de classes [63], situation souvent appelée non-exclusivité. Notons à ce stade que certaines méthodes itératives (les méthodes markoviennes, par exemple) se basant sur l'utilisation de niveaux de gris, exploitent le contexte spatial des pixels afin de remédier au problème de chevauchement spectral. Néanmoins, ces méthodes propagent souvent une information de nature imprécise, ce qui augmente l'ambiguïté lors de la classification pixelique [111].

La théorie des ensembles flous, introduite par L. Zadeh [60], fournit un cadre conceptuel pour la représentation des connaissances ambiguës. Généralement, l'application de cette théorie dans le processus d'analyse de scène consiste à associer à chaque objet observé (pixel, région ou objet) de l'image, c'est-à-dire à chaque valeur numérique mesurée de la primitive caractérisant l'objet, des degrés d'appartenance (dans l'intervalle $[0, 1]$) aux différentes classes. Ces degrés d'appartenance sont ensuite utilisés dans le calcul d'une mesure floue (l'entropie floue [85], par exemple). Une telle mesure permet d'attribuer une « décision finale » concernant l'appartenance de l'objet à l'une des classes considérées.

En effet, l'utilisation d'un tel cadre mathématique pour représenter et intégrer les connaissances ambiguës, rend la conception de méthodes de reconnaissance de formes, basée sur des modèles interprétables, plus facile et capable de résoudre les problèmes d'ambiguïté rencontrés lors d'un processus d'interprétation [26].

Parmi les méthodes basées sur la théorie des ensembles flous, l'algorithme de Fuzzy C-Means et les systèmes à base de règles floues sont très utilisés dans la littérature.

1.4.4.1 L'algorithme Fuzzy C-Means (FCM) :

L'algorithme FCM [94] fournit un moyen classique pour résoudre le problème de regroupement ou « *clustering* » selon les degrés d'appartenance binaires (soit 0, soit 1) de chaque objet à une classe. Il permet d'attribuer à chaque objet des valeurs d'appartenance à chacune des classes thématiques dans l'intervalle $[0, 1]$.

Pour une image comportant M classes thématiques (connues *a priori*), le principe de base de cet algorithme consiste à représenter chaque classe j par un vecteur de référence v_j , $j=1, \dots, M$. On suppose disponible à l'initialisation, un ensemble de N vecteurs de mesures x_i , $i=1, \dots, N$ de même dimension, chaque vecteur x_i correspondant à un objet i . Pour chacun de ces vecteurs x_i , on considère également les degrés d'appartenance μ_{ij} de ce vecteur x_i à la classe j . L'algorithme FCM a alors pour objectif d'affiner de façon itérative, les μ_{ij} et les M vecteurs de référence v_j , $j=1, \dots, M$, en minimisant la fonction objective suivante :

$$J(U, V) = \sum_{i=1}^N \sum_{j=1}^M \mu_{ij} (x_i - v_j)^2 \quad (1.5)$$

avec $V = \{v_1, v_2, \dots, v_M\}$ et U la matrice $[\mu_{ij}] \in [0, 1]$, $i=1, \dots, N$, $j=1, \dots, M$ telle que $\sum_{j=1}^M \mu_{ij} = 1$ représentant la partition floue générée par l'algorithme. La distance Euclidienne $(x_i - v_j)^2$ peut être remplacée par la distance de Mahalanobis, la matrice U devenant alors une matrice de covariance floue (plus de détails sont donnés dans le chapitre 3).

1.4.4.2 Les systèmes à base de règles floues:

Une règle floue est une construction qui a la forme générale « *si A alors B* », où A et B sont des propositions contenant des termes linguistiques. En effet, l'utilisation des termes linguistiques dans ces

règles permet d'une façon flexible, de représenter les connaissances humaines et d'exprimer l'imprécision et l'incertitude inhérentes dans ces connaissances. Chaque règle floue possède un niveau de certitude, et ceci forme la différence entre une règle floue et une règle classique qui utilise la logique binaire [63].

Généralement, un système à base de règles floues est composé de trois étapes principales (Figure 1.4). La première étape est la « *fuzzification* » qui implique la partition de l'espace des primitives « *entrées* » en sous-espaces flous, chacun étant caractérisé par une fonction d'appartenance (voir chapitre 3). Les règles floues sont ensuite générées pour chaque sous-espace. La deuxième étape est l'« *inférence* » qui nécessite le calcul du niveau de certitude de chaque règle floue et de combiner les différentes règles floues. L'étape finale est la « *defuzzification* » : elle consiste à transformer la sortie floue des règles de décision en des « *sorties* » binaires pour le système [63].

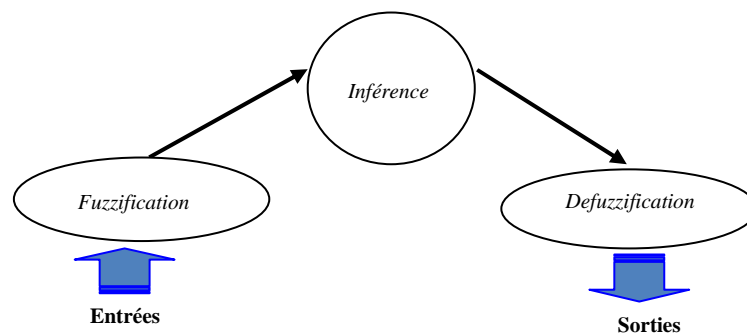


Figure 1.4 : Architecture générale d'un système à base de règles floues

1.4.5 Les méthodes de fusion d'information multi-sources

Généralement, il est très difficile pour une méthode de reconnaissance de formes d'aborder tous les aspects du problème traité (discrimination entre les classes, traitement d'imperfections, etc.) en se basant sur une seule source de connaissance. En outre, une méthode de reconnaissance de formes peut être meilleure que d'autres méthodes pour résoudre un aspect particulier du problème traité, tandis que d'autres méthodes peuvent être supérieures sur un autre plan [63]. Ceci a motivé, en reconnaissance de formes, l'utilisation conjointe (ou la fusion) de plusieurs sources de connaissances hétérogènes. Cette utilisation permet à une méthode de reconnaissance de formes d'exploiter la complémentarité et la redondance de ces sources de connaissances afin de réaliser des inférences qui ne sont pas réalisables à partir d'une seule source (par exemple, réduire l'incertitude dans un processus de classification).

Dans le domaine de la vision par ordinateur, le processus d'analyse de scène est fortement lié à la problématique de fusion de plusieurs sources des connaissances [10]. Ceci est due à l'émergence de plusieurs types de capteurs d'image (visible, infrarouge, radar, etc.) fournissant différentes mesures physiques (images multi-résolution, images multi-temporelles, images multi-bandes spectrales et images multi-polarisations). Notons que d'autres sources de connaissances peuvent être obtenues à partir des calculs effectués sur les mesures numériques fournies par un seul capteur ou des sources de connaissances liées à l'expert [71]. En effet, la fusion d'images issues de plusieurs sources de connaissances consiste à les combiner en une seule image « fusionnée » ayant un contenu « amélioré » rendant ainsi la mission de l'expert, plus facile pour détecter, reconnaître et identifier les objets présents dans la scène [28]. Par conséquent, le processus d'interprétation de scène peut être réalisé en utilisant des méthodes de reconnaissance de formes basées sur la fusion d'informations issues de telles sources [29]. Toutefois, en analyse de scène, il y a plusieurs niveaux de reconnaissance : le niveau capteurs, le niveau des primitives et le niveau des décisions (Figure 1.5) et à chaque niveau, nous pouvons identifier

différents types d'imperfections [46]. Par conséquent, le problème de la gestion de ces imperfections devient de plus en plus complexe en fusion d'informations [9] et les systèmes d'interprétation doivent être en mesure de faire face à ce type d'information [1].

Le processus de fusion est composé principalement de quatre étapes : 1) la modélisation, 2) l'estimation, 3) la combinaison et 4) la décision [29]. L'étape de modélisation comprend le choix d'une méthode de représentation des connaissances adaptée à l'approche de fusion retenue. L'étape d'estimation consiste alors à déterminer les distributions numériques nécessaires pour estimer les informations à fusionner. La troisième étape consiste à choisir un opérateur approprié pour la fusion. L'étape finale de décision implique le choix d'une règle de décision basée sur les informations fournies par les différentes sources.

D'une manière générale, le processus de fusion en interprétation de scènes observées par des capteurs d'images peut être réalisé selon quatre niveaux différents [71] (Figure 1.5) :

- Fusion au niveau "signal" : les signaux provenant de différents capteurs sont combinés pour créer un nouveau signal avec un rapport signal sur bruit meilleur que le rapport des signaux d'origine, c'est-à-dire avant la fusion.
- Fusion au niveau "pixel" : cette fusion est effectuée pixel par pixel. Chaque pixel de l'image fusionnée est déterminé à partir d'un ensemble de pixels correspondants dans les images sources.
- Fusion au niveau "primitive" : cette fusion nécessite une étape d'extraction de primitives à partir d'un ensemble de pixels (intensités, contours, zones de textures,). Les primitives de même nature sont ainsi combinées.
- Fusion au niveau "décision" : cette fusion opère dans un niveau d'abstraction plus élevé que les précédents niveaux et combine les résultats de plusieurs algorithmes pour aboutir à une décision finale fusionnée. Les images d'entrée sont traitées individuellement pour l'extraction d'informations. Les informations obtenues sont ensuite combinées en appliquant des règles de décision pour avoir une décision finale.

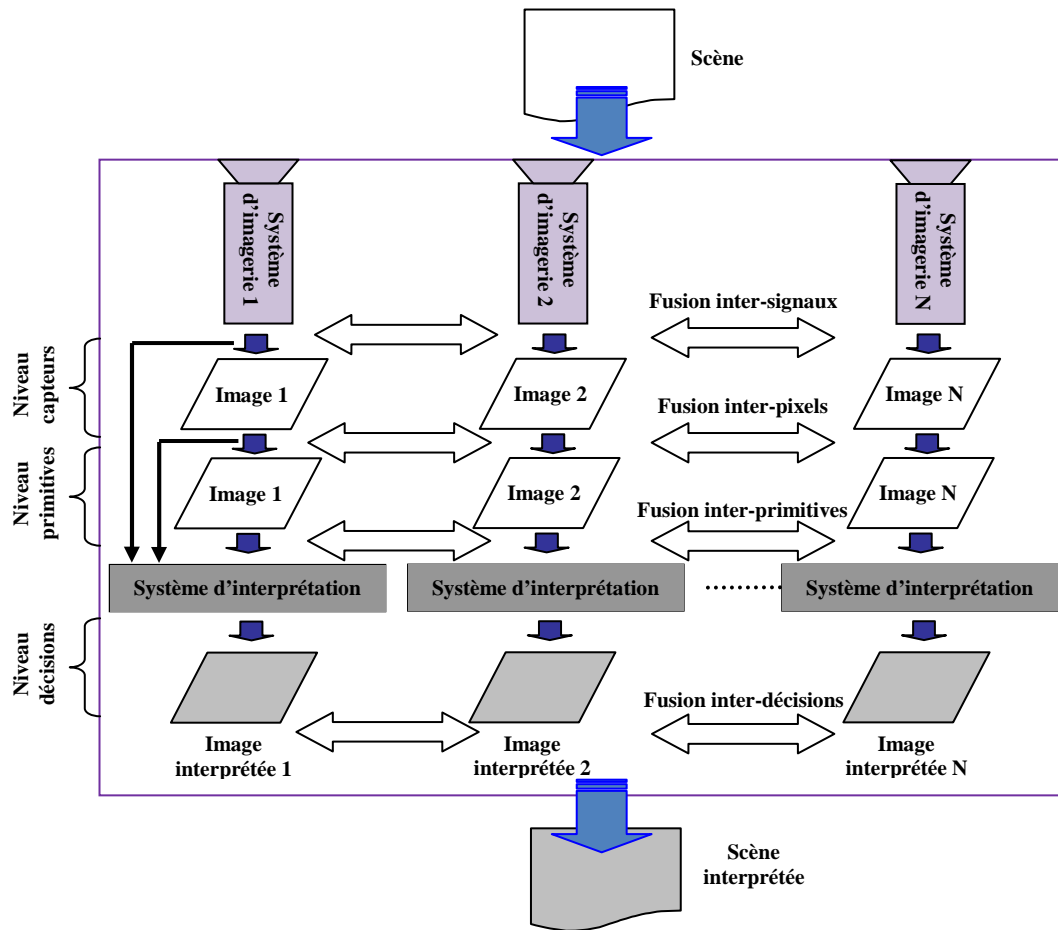


Figure 1.5 : Les différents niveaux de fusion des données multi-capteurs

La théorie des probabilités, la théorie des croyances, la théorie des ensembles flous et la théorie des possibilités constituent les principaux outils mathématiques les plus rencontrés dans la littérature pour fusionner les informations issues de plusieurs sources de connaissances. Les techniques de fusion associées à la théorie des possibilités seront plus spécifiquement présentées dans le chapitre 2 car cette théorie constitue le cadre de notre travail de recherche.

1.5 Niveaux d'interprétation d'une scène

Au cours des trois dernières décennies, plusieurs systèmes d'interprétation de scènes ont été mis en œuvre dans différents domaines d'application [1]. Ces systèmes ont été conçus selon différentes philosophies, modèles et motivations. Ils diffèrent par un certain nombre d'aspects tels que l'objectif final, le niveau des connaissances disponibles et la représentation de ces connaissances, la méthodologie à suivre et le domaine d'application. Généralement, la méthodologie utilisée par ces systèmes sert de critère pour classer ces systèmes d'interprétation de scène en différentes familles [1] [3]: les systèmes basés sur les connaissances [30], les systèmes basés sur une théorie de décision particulière [31], les systèmes basés sur le mode de raisonnement [32], etc.

La sémantique dans une image dépend d'un côté, du niveau des connaissances manipulées et de la perception qu'a l'expert de cette image et d'un autre côté, de l'objectif poursuivi par l'expert [5]. Cet objectif détermine la nature du processus d'interprétation et l'échelle de la zone d'étude, ce qui définit ainsi les résolutions spatiale et spectrale convenables. Ainsi, partant de ce constat, nous proposons de classer les systèmes d'interprétation de scène en fonction des niveaux sémantiques sur lesquels l'interprétation est opérée. Ces niveaux sont :

- Le niveau sous-pixelique,
- Le niveau pixelique,
- Le niveau région,
- Le niveau objet,
- Le niveau scène.

Chaque niveau sémantique est caractérisé par les connaissances nécessaires et le mode de représentation pour mettre en œuvre les objectifs à atteindre. Plus le niveau sémantique devient abstrait, plus des connaissances *a priori* sont indispensables et plus les modes de représentation deviennent complexes et de ce fait, la gestion des imperfections devient cruciale.

Dans ce paragraphe, nous allons décrire chaque niveau sémantique d'interprétation, la représentation des connaissances associées et montrer des exemples de systèmes d'interprétation opérant sur ces niveaux. Un intérêt particulier est porté sur les trois premiers niveaux (sous-pixelique, pixelique, région) en explicitant les problèmes rencontrés tels que l'intervention de l'expert lors du processus d'interprétation (i.e. les paramètres à estimer), la représentativité des connaissances disponibles, l'exploitation du contexte spatial comme une source additionnelle de connaissances, la possibilité de mise en cause par l'expert des résultats obtenus aux niveaux supérieurs (région, objet et scène) et le traitement des imperfections.

1.5.1 Le niveau sous-pixelique

Ce niveau d'interprétation de scène concerne l'identification du contenu informationnel des pixels comportant un mélange de plus d'une classe parmi les différentes classes thématiques contenues dans la scène analysée (pixels mixtes). Ce phénomène, qui correspond au fossé sensoriel [3], est dû essentiellement à la résolution spatiale des capteurs et au chevauchement spectral entre les informations des classes [63]. Dans le domaine spectral, les caractéristiques de réflectance, émission, ou rétrodiffusion des différentes classes peuvent être similaires. Dans le domaine spatial, un pixel, une région, ou même un objet peut contenir des zones représentatives de plus d'une seule classe.

Les connaissances *a priori* concernant ce niveau sont d'une part, liées aux mécanismes d'interaction entre le paramètre physique mesuré et les différentes substances ou classes pures existantes dans la scène (souvent appelées *endmembers*) et d'autre part, aux résolutions spatiale et spectrale du système d'imagerie. Ces connaissances permettent la sélection convenable tant des données disponibles que des méthodes de traitement afin d'identifier le contenu sous-pixelique.

Les caractéristiques spectrales d'une classe pure sont souvent représentées par un spectre unique et peuvent être extraites automatiquement à l'aide d'algorithmes non supervisés [63] ou supervisés lorsque les données d'apprentissage sont disponibles [39].

Deux approches sont principalement utilisées pour modéliser la contribution des différentes classes pures au spectre mesuré à partir d'un pixel : le *modèle linéaire* qui considère que ce spectre mixte est la somme pondérée des spectres des classes pures et le *modèle non-linéaire* qui tient compte de la diffusion multiple du rayonnement incident (solaire, onde radar, onde ultrason, etc.) [13] (voir chapitre 4 pour plus de détails). Ainsi, le processus de démixage « *unmixing* » des pixels mixtes vise à identifier les classes pures constituant le spectre mesuré d'un pixel mixte ainsi que la proportion de chacune de ces classes présentes dans le pixel analysé (souvent appelée *taux d'abondance*). D'une manière générale, le processus de démixage est réalisé en utilisant la notion d'appartenances multiples et partielles de chaque pixel (représentation floue [12]) et dans chaque bande spectrale, aux différentes classes pures.

S. Verbeiren *et al* [40] ont proposé une approche pour la classification sous-pixelique des zones de cultures régionales. Cette approche est composée de deux étapes. Dans un premier temps, le modèle linéaire est adopté pour générer des cartes d'abondance à partir d'images SPOT. Ensuite, un réseau de neurones est utilisé pour leur classification. L'apprentissage du réseau de neurones est effectué sur la base d'une vérité terrain validée par un expert. La construction d'une base d'apprentissage représentative de la scène observée ainsi que le non-traitement d'imperfections entachant les données extraites forment les principales limitations de cette approche.

Z. B. Rabah [3] a proposé une approche pour l'interprétation des données de type hyper-spectrale (IHS). Cette approche a permis dans un premier temps, d'effectuer une analyse spectrale basée sur un modèle de mélange linéaire et d'une mesure de similarité spectrale (angle spectral) entre la signature spectrale calculée pour chaque pixel et une base de signatures spectrales de référence. L'objectif de cette phase est de retrouver la composition quantitative et qualitative de chaque pixel de l'image (cartes d'abondance). Dans un deuxième temps, les connaissances extraites sont interprétées en les fusionnant avec d'autres sources de connaissances liées à l'expert pour aboutir à une cartographie précise des phénomènes à l'origine des risques naturels menaçant la terre. Cette phase d'interprétation a été réalisée par une fusion possibiliste basée sur un modèle neuro-flou qui permet de traiter conjointement les problèmes d'ambiguïté entachant les IHS et d'incertitude caractérisant les méthodes de démixage spectral. Un aspect très important de cette approche est liée à l'exploitation conjointe de plusieurs sources de connaissances (connaissances extraites à partir des données disponibles ainsi que les connaissances d'expert). Cependant, cette approche nécessite une intervention active de l'expert dans toutes les étapes : la construction d'une base de règles floues, la définition des paramètres du modèle neuro-flou ainsi que la constitution d'une base d'apprentissage représentative.

D'une manière générale, nous constatons que les approches d'interprétation de scène au niveau sous-pixelique, sont basées essentiellement sur l'utilisation de multiples images (multi-spectrales ou hyper-spectrales) et adoptent le modèle linéaire pour réaliser le démixage. Néanmoins, ces approches abordent le problème des pixels mixtes seulement « pixel par pixel » ce qui rend très compliqué l'identification du contenu informationnel de régions ou d'objets comportant un mélange de plus d'une classe parmi les différentes classes pures contenues dans la scène.

Ainsi, à ce stade, la question qui se pose est : "comment peut-on réduire l'ambiguïté liée à l'existence des pixels mixtes, dans le cas où on ne dispose que d'une seule source d'image et ceci dans un contexte de haute résolution spatiale ?". Généralement dans ce contexte, les méthodes utilisées dans le processus d'interprétation de scène sont des méthodes de reconnaissance de formes opérant soit au niveau pixelique ou à un niveau plus élevé (région, objet, scène) et donc ne prennent pas compte le contenu informationnel sous-pixelique.

1.5.2 Le niveau pixelique

A ce niveau, les systèmes d'interprétation de scène considèrent le pixel comme la granularité informationnelle de base dans le processus d'interprétation, puisque ce pixel résulte d'une empreinte correspondant à une cellule de résolution du capteur. L'objectif consiste alors à identifier le contenu informationnel de chaque pixel individuellement. Ce contenu est associé soit à la classe unique du pixel (parmi les différentes classes de la scène), soit à un élément d'information caractérisant chaque pixel : information de détection de changement, des paramètres physiques (taux d'humidité, rugosité, etc.), des paramètres 3D permettant la construction d'un modèle numérique de terrain,...

Au niveau pixelique, les connaissances *a priori* sont basées soit sur les informations du niveau sous-pixelique (l'appartenance de chaque pixel de l'image aux différentes classes thématiques dans la scène)

soit sur une représentation statistique extraite des mesures physiques en sortie du capteur, et qui caractérise la fréquence d'occurrence des valeurs mesurées pour chaque pixel [63].

Le mode de représentation probabiliste est effectué au moyen des distributions de probabilités sur l'ensemble des classes où chaque pixel (dans chaque espace de primitive) est associé à une valeur de probabilité caractérisant l'occurrence de ce pixel dans la classe considérée. Ce mode de représentation probabiliste permet de modéliser l'incertitude inhérente aux connaissances disponibles.

Les méthodes bayésiennes, basées sur la représentation probabiliste, sont utilisées dans plusieurs travaux d'interprétation de scène au niveau pixelique.

L. Bruzzone et D. F. Prieto [33] ont proposé une approche pour la détection de changement automatique. La technique proposée est basée sur l'hypothèse que la densité de probabilité $Pr(x)$ calculée sur les valeurs de niveau de gris de pixels dans l'image de différence, peut être modélisée comme une densité d'un mélange de deux classes : classe de pixels non-modifiés et classe de pixels modifiés. Ces deux classes sont supposées avoir des densités de probabilité gaussiennes dont les paramètres sont estimés par l'approche de maximisation d'espérance (EM).

A. Villa *et al.* [34] ont utilisé l'analyse discriminante en composantes indépendantes « *ICDA : Independent Component Discriminant Analysis* » pour la classification pixelique en télédétection. Les fonctions de densité de probabilités de chaque composante indépendante sont estimées en utilisant la méthode d'estimation par noyau « *KDE : Kernel Density Estimation* » et enfin, la règle de Bayes est appliquée pour la décision de classification finale.

L'algorithme de maximum de vraisemblance (ML) a été largement utilisé pour la classification thématique des images surtout dans le domaine de la télédétection [35][63]. M. Törmä [43] a essayé d'améliorer le taux de reconnaissance dans les zones naturelles en utilisant une méthode de fusion multi-sources. Les images ont été d'abord classifiées en utilisant l'algorithme de maximum de vraisemblance et les résultats de classification ont été fusionnés à l'aide de différentes méthodes de fusion basées sur le théorème de Bayes.

L'appartenance d'un pixel à une classe donnée dépend de l'occurrence statistique de la valeur numérique mesurée pour la classe considérée. Ceci rend les méthodes basées sur la représentation probabiliste très sensibles au bruit et incapables de discriminer les classes dans un contexte où ces classes sont très difficiles à discriminer (c'est-à-dire avec la présence d'un fort chevauchement des valeurs observées pour les différentes classes). De plus, la représentation probabiliste ne permet pas d'intégrer d'une façon flexible, les connaissances de type descriptif de l'expert qui sont, par nature, ambiguës et liées au contenu sémantique de l'image.

Afin de remédier à ces inconvénients liés à la représentation probabiliste, d'autres modes de représentation des connaissances ont été utilisés.

F. Ludic [36] a développé un outil de classification pixelique d'images « *Full : Fuzzy Reasoning applied to Image Intelligence* ». L'outil *Full*, implémenté dans une boîte à outils ENVI, est basé sur la représentation des connaissances selon les théories des ensembles flous et des croyances de Dempster-Shafer. La représentation floue des connaissances est obtenue au travers de fonctions d'appartenance aux formes prédéfinies : gaussienne, triangulaire, trapézoïdale ou la forme de l'histogramme normalisé. Les paramètres de ces fonctions sont estimés à partir des échantillons sélectionnés par l'expert. Les fonctions de masse sont ensuite calculées à partir des fonctions d'appartenance en utilisant des méthodes de transformations telles que la normalisation des valeurs d'appartenance. Le but de cet outil est de comparer plusieurs opérateurs de fusion et règles de décision dans le cadre de la classification d'images appliquée à la cartographie de la couverture du sol. La sélection par l'expert, des échantillons considérés

comme totalement représentatifs des différentes classes dans la scène et ensuite l'utilisation de ces échantillons pour l'estimation des paramètres des fonctions d'appartenance constituent l'inconvénient majeur de cette méthode.

L. Roux and J. Desachy [37] ont présenté une méthode de fusion d'informations multi-sources pour la classification d'images satellitales. Les principales caractéristiques de cette méthode sont l'utilisation de la théorie des possibilités pour gérer l'incertitude liée à la classification des pixels et la possibilité de fusionner les sources numériques (les bandes spectrales de l'image). Les distributions de possibilités sont simplement estimées par la normalisation d'histogrammes des données d'apprentissage et ceci dans chaque bande spectrale. Cette méthode a été appliquée sur des images du satellite Landsat avec un taux de reconnaissance de l'ordre de 78%. Ce résultat, très modeste, est dû à la négligence du contexte spatial des pixels à classifier ainsi que la forte dépendance statistique des données d'apprentissage sélectionnées par l'expert.

T. Nakashima *et al* [83] ont proposé une méthode de classification basée sur des règles floues qui permettent de mettre l'accent sur une ou plusieurs classes. Ils ont considéré le problème de classification comme un problème de minimisation des coûts et ont employé la notion de pondération dans le processus d'apprentissage. Les règles floues sont générées en tenant compte des pondérations des données d'apprentissage, et en utilisant des ensembles flous triangulaires. Cette méthode a été évaluée sur plusieurs applications de traitement d'image, telles que le diagnostic du cancer du sein à partir d'images mammographiques. Le taux de reconnaissances était de l'ordre de 94% (dans le cas d'images mammographiques) et de l'ordre de 64% (dans le cas de la classification des images du satellite Landsat, base de données *STATLOG*). L'intervention importante de l'expert tant au niveau de la construction des règles floues qu'au niveau de pondération constitue un inconvénient essentiel de cette méthode. De plus, la négligence du contexte spatiale des pixels ne permet pas la bonne discrimination des classes ayant un fort chevauchement (comme c'est le cas pour les données *STATLOG*).

Farah *et al* [29] ont proposé une approche semi-automatique pour l'interprétation des images, en tenant compte des imperfections qui accompagnent ce processus. Cette approche comporte trois modules. Le premier module est un module d'apprentissage basé sur l'évaluation de trois méthodes de fusion (une méthode probabiliste, une méthode possibiliste et une méthode reposant sur la théorie des croyances) appliquées aux images satellitales. Le deuxième module vise à déterminer la meilleure méthode de fusion en utilisant une base de règles et un raisonnement à base de cas (CBR). Le dernier module est consacré à la fusion d'images multi-capteurs en utilisant la meilleure méthode de fusion récupérée par le CBR. Les distributions des probabilités (respectivement, des possibilités et celles des masses de croyance) sont estimées par la méthode de l'estimation à base de noyau « *KDE : Kernel Density Estimation* » [107] (respectivement, par l'algorithme de possibilistic C-means (PCM) [109]) à partir des régions d'intérêt choisies par l'expert. Cette approche a été validée sur un ensemble d'images optiques SPOT-4 et des images radar ERS-2 représentant une région tunisienne centrale. Nous constatons dans cette approche l'utilisation d'une méthode non-paramétrique *KDE* pour l'estimation des distributions de probabilités à partir d'un ensemble limité d'échantillons sélectionnés par l'expert. Cependant, cette approche proposée dans [29] présente des limitations liées à la construction d'une base des règles ainsi qu'à la dépendance des régions d'intérêt choisies par l'expert pour l'estimation des distributions (probabiliste, possibiliste et celle des croyances) caractérisant les différentes classes dans la scène. Le contexte spatial est aussi négligé dans cette méthode d'interprétation de scène.

D'une manière générale, l'utilisation des représentations floue et possibiliste dans les méthodes proposées pour l'identification du contenu informationnel et l'interprétation de scène au niveau pixelique permet, d'un côté de gérer l'attribution des pixels aux différentes classes par rapport à la représentation

probabiliste et d'un autre côté, d'intégrer les connaissances d'expert liées au contenu sémantique de l'image. En revanche, ces méthodes opèrent sur des ensembles limités d'échantillons d'apprentissage et qui sont considérés comme représentatifs des différentes classes dans la scène. De plus, le contexte spatial des pixels permettant de réduire l'incertitude liée à la classification de ces pixels n'est pas exploité dans la majorité de ces méthodes. En effet, ces méthodes ne fournissent pas les mécanismes qui permettent d'évaluer l'incertitude liée à la classification des pixels.

1.5.3 Le niveau région

Ce niveau d'interprétation porte sur l'identification du contenu informationnel des entités homogènes (régions) formées par des pixels connectés qui ont des caractéristiques similaires telles que les caractéristiques colorimétriques, texturales ou morphologiques. Généralement, le processus d'interprétation de scène au niveau région commence par une phase d'analyse basée sur des techniques de segmentation des images disponibles afin d'extraire les régions homogènes de la scène.

Les connaissances *a priori* liées aux caractéristiques des régions permettent généralement d'exprimer l'apparence visuelle de chaque classe de régions. Cela peut se traduire par des termes linguistiques ou des descripteurs qualitatifs qui peuvent être représentés par des fonctions d'appartenance ou par la délimitation des échantillons considérés comme représentatifs des régions à identifier. Ces échantillons servent soit à estimer des distributions (probabiliste, d'appartenance, possibiliste ou des masses de croyance) caractérisant les régions d'intérêt [29], soit à extraire des caractéristiques texturales telles que des coefficients d'homogénéité basés sur la matrice de cooccurrence (GLCM)) [27], etc.

F. Liu *et al* [41] ont proposé une méthode de segmentation d'images pour l'identification du contenu des plaques d'athérosclérose dans l'être vivant. Cette méthode est composée de deux parties. Dans la première partie, des cartes de probabilités sont générées en se basant sur des réseaux bayésiens. A chaque pixel de ces cartes, une valeur de probabilité d'appartenance à une classe donnée est attribuée individuellement. Les densités de probabilités conditionnelles sont estimées à partir d'un ensemble limité d'échantillons d'apprentissage par le biais d'une estimation basée sur la méthode à base de noyau « *KDE : Kernel Density Estimation* ». Dans la seconde partie, ces cartes de probabilités et un modèle basé sur les champs de Markov (MRF) sont combinés afin d'exploiter une règle de décision basée sur le maximum *a posteriori* (MAP) et ceci afin d'attribuer une étiquette de classe à chaque région homogène. Dans cette méthode, nous constatons l'utilisation de la méthode *KDE* pour l'estimation des distributions de probabilités à partir d'un ensemble limité d'échantillons d'apprentissage ainsi que l'exploitation du contexte spatial afin de réduire l'incertitude liée à la classification des régions détectées. L'estimation à partir d'un ensemble limité d'échantillons constitue une contrainte de cette méthode. De plus, les résultats obtenus par l'utilisation des champs de Markov dépendent fortement des conditions initiales [42].

F. Huet et S. Philipp [100] ont aussi proposé un système de fusion d'images. Dans un premier temps, l'image analysée est segmentée par plusieurs méthodes de segmentation. Ensuite, chaque région détectée dans les images segmentées est interprétée en lui attribuant une valeur d'appartenance à une classe donnée. Ces valeurs d'appartenances sont estimées à partir des données d'apprentissage et en utilisant l'algorithme des K plus proches voisins K-NN [99]. Les images interprétées sont ensuite fusionnées par un opérateur de fusion de la théorie des ensembles flous. Plusieurs opérateurs de fusion ont été comparés. Cette méthode présente l'intérêt d'utiliser la représentation floue ainsi que la fusion multi-sources des connaissances (plusieurs images segmentées) afin d'améliorer les résultats de la segmentation. Cependant, cette méthode dépend d'un ensemble limité de d'échantillons d'apprentissage pour

l'estimation des valeurs d'appartenance des régions détectées. De plus, aucune évaluation de l'incertitude liée à la classification de ces régions n'a été réalisée dans ce système.

W. Eziddin [111] a présenté une approche de segmentation d'images par diffusion des connaissances dans le domaine possibiliste, approche qui a pour objectif de détecter des tumeurs en imagerie mammographique. Un aspect important du raisonnement humain a été modélisé dans ce processus de classification des pixels en intégrant de manière itérative, l'information issue du contexte spatial de ces pixels. En outre, cette approche permet d'évaluer les valeurs d'appartenance des pixels afin d'affiner les résultats de segmentation. Cette approche peut être essentiellement résumée en quatre étapes (Figure 1.6):

- L'initialisation de l'approche qui est basée sur des connaissances descriptives exprimées par l'expert. Ces connaissances sont représentées par des distributions de possibilités ayant des formes prédéfinies et associées aux différentes classes thématiques présentes dans l'image. La projection de ces connaissances sur l'image initiale permet de générer des images appelées « cartes possibilistes » représentant l'appartenance de chaque pixel de l'image initiale aux différentes classes thématiques.
- Le système de décision qui a pour objectif d'attribuer les pixels ayant une appartenance à une classe thématique donnée avec un fort degré de certitude, à cette classe, et ceci selon la règle de décision adoptée. Ces pixels classifiés avec une forte certitude sont appelés *germes possibilistes* et forment les connaissances extraites à chaque itération. Les pixels qui ne sont pas classés durant cette itération sont considérés comme appartenant à une classe nommée « rejet ». Cette classe est une classe de rejet de classification temporaire; en effet, l'appartenance aux différentes classes, des pixels qui y sont rangés, sera réévaluée durant l'itération suivante.
- Le système de diffusion des connaissances possibilistes se charge ensuite de diffuser itérativement dans le voisinage des pixels les connaissances extraites de l'image (i.e. germes possibilistes), dans l'espace des connaissances possibilistes. Cette étape aboutit à la mise à jour des cartes possibilistes qui seront utilisées à l'itération suivante. Deux modes d'intégration des germes possibilistes (i.e. point d'encrage des classes) ont été proposés : statique et dynamique. De même, en se basant sur le principe du filtre de Nagao [117], deux stratégies de diffusion ont été proposées : la méthode de diffusion classe/classe et la diffusion croisée. Dans la première stratégie, le choix de la clique de diffusion optimale aboutit à la sélection de différentes cliques en fonction de la carte possibiliste considérée. En revanche, pour la stratégie de diffusion croisée, une même clique est appliquée pour la diffusion des connaissances sur toutes les cartes. Ces deux stratégies de diffusion des connaissances, ainsi que les deux modes d'intégration des germes ont été évaluées de manière qualitative et quantitative.
- Le système d'évaluation qui permet, d'un côté d'évaluer l'incertitude liée à la classification des pixels afin d'enrichir l'ensemble des germes possibilistes et d'un autre côté, d'évaluer la nécessité de poursuivre le processus itératif de diffusion possibiliste des connaissances.

La nature itérative de cette approche permet alors de propager les connaissances extraites à toutes les zones de forte ambiguïté de l'image. La prise de décision de classer un pixel, retardée jusqu'au moment où la connaissance propagée sera « jugée » suffisante, permet également de réduire l'incertitude inhérente à la classification des pixels. Les résultats obtenus par cette approche ont montré leur supériorité par rapport aux approches classiques (l'approche LevelSet [113], l'approche FCM [97] et l'approche de segmentation par croissance de régions [114]). Cependant, l'estimation initiale des distributions de possibilités nécessite l'intervention de l'expert afin de sélectionner les formes de base de

ces distributions ainsi que le besoin d'appliquer des algorithmes d'optimisation pour le réglage automatique des paramètres associés à ces distributions.

Notons que ce travail a constitué le point de départ d'une partie de la méthodologie d'interprétation de scène que nous proposons dans notre travail. Cette partie concerne le raffinement et la segmentation par diffusion des connaissances possibilistes.

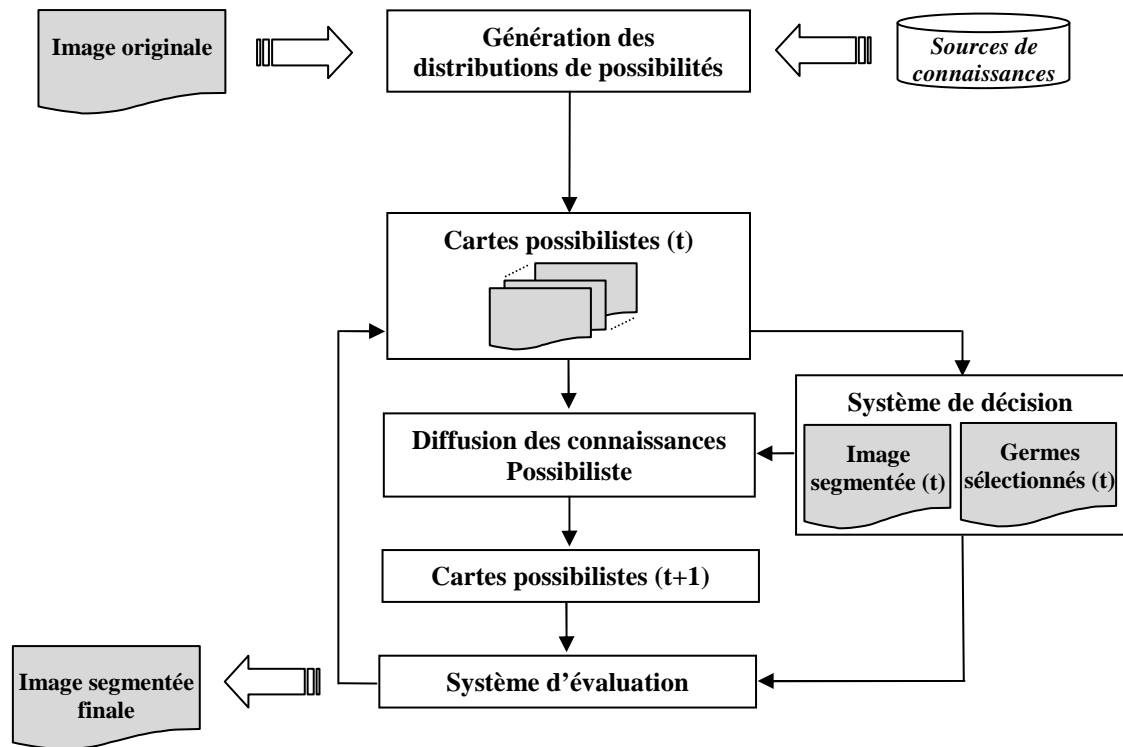


Figure 1.6 : Méthode de segmentation d'image par diffusion des connaissances possibilistes

Les méthodes d'interprétation proposées au niveau région, ont de plus en plus tendance à utiliser les représentations floue ou possibiliste dans le processus d'interprétation de scène et ceci dans un contexte de plusieurs sources des connaissances : contexte spatial des pixels ou des régions, connaissances d'expert, résultats de plusieurs méthodes de segmentation, etc. Cette tendance est due au besoin croissant d'intégrer des connaissances *a priori* d'une part, et de manipuler explicitement les imperfections liées aux connaissances d'autre part.

Notons que d'une manière générale, les méthodes proposées dans la littérature suivent une démarche « ascendante » du niveau sémantique le plus bas (niveau sous-pixelique, ou pixelique) vers le niveau objet. Aucune méthode ne propose une démarche descendante permettant de déterminer, pour les régions détectées, les caractéristiques fondamentales du type composition/contenu global des classes thématiques de base. Nous proposons d'investiguer une telle démarche descendante d'interprétation dans le cadre du système d'interprétation présenté dans ce travail.

1.5.4 Le niveau objet

La présence d'une ou de plusieurs entités du type région (ayant des structures et des caractéristiques spécifiques) peut aboutir à la constitution d'un contenu informationnel d'un niveau sémantique plus élevé que celui possédé par ces régions. L'ensemble de ces entités "région" forment un « objet » et les

connaissances *a priori* nécessaires à ce niveau d'interprétation de scène sont liées aux différentes structures et caractéristiques des régions ainsi qu'à la description sémantique des objets eux-mêmes (connaissances qualitatives ou de haut niveau) [14]. La présence conjointe de régions ayant une forme rectangulaire et de régions d'ombres peut suggérer l'existence de bâtiments ; la présence conjointe d'yeux et d'une bouche peut caractériser l'occurrence d'un visage (objets 3D), etc.

Les connaissances qualitatives décrivant les objets concernent généralement, les aspects géométriques (forme, localisation, orientation, etc.), morphologiques, radiométriques et contextuels des régions qui forment l'objet. La représentation des connaissances *a priori* permet, au niveau « objet », de s'appuyer sur des stratégies de raisonnement pour combiner des primitives simples ou régions extraites par les niveaux inférieurs, au sein de primitives plus complexes formant des objets. Cette représentation peut être faite par plusieurs méthodes. Par exemple, la méthode « frame » qui est un schéma contenant toutes les connaissances *a priori* liées à un type donné d'objets [16] constitue un exemple type.

Plusieurs besoins de l'expert peuvent s'exprimer à ce niveau tel que la classification d'objet, le suivi d'objet, la vidéo-surveillance, la reconnaissance d'objets 3D, etc. V. Amberg [15] a proposé une approche hiérarchique d'interprétation de scènes à partir d'une unique image complexe d'un radar à ouverture de synthèse haute résolution. Cette approche a été appliquée pour l'extraction quasi-automatique du réseau routier. La stratégie d'interprétation de scène consiste à travailler à plusieurs niveaux en détectant à chaque niveau des structures ou objets de plus en plus petits. Une étape de classification de scène a été réalisée dans un premier temps afin de formuler une idée rapide de l'occupation des sols d'une scène et piloter ainsi le processus suivant d'extraction. Dans cette étape, une méthode de classification bayésienne contextuelle a été proposée. Cette méthode inclut un algorithme de classification ponctuelle au sens du maximum de vraisemblance d'une image prétraitée afin d'éliminer le bruit du speckle et une règle de décision majoritaire. Dans un deuxième temps, une méthode d'extraction des réseaux routiers a été proposée. Cette méthode comporte deux étapes : 1) la localisation des tronçons rectilignes des routes de la scène en utilisant la transformation de « Hough » et 2) l'extraction du tracé des routes par un algorithme de suivi. Cette approche présente une analyse multi-échelle permettant de focaliser l'analyse sur les objets « évidents » que l'œil humain peut détecter facilement puis d'exploiter les connaissances liées à ces objets afin de guider plus finement la recherche des autres objets. Certains des algorithmes proposés lors de cette approche, en particulier l'algorithme de suivi, présentent un nombre non négligeable de paramètres à choisir. Ceci limite fortement les performances de cette approche.

F. Prandi *et al* [43] ont proposé une approche semi-automatique basée sur des règles floues pour l'extraction et la reconnaissance des objets 3D dans les zones urbaines. Dans un premier temps, les composants structuraux descriptifs (connaissances 3D ou géométriques) de l'objet sont extraits à partir d'une adaptation numérique de surface « *DSM : Digital Surface Matching* ». Les DSM ont été utilisées afin d'automatiser et d'accélérer la reconnaissance des objets topographiques dans des scènes complexes. Ensuite, les composants extraits ont été intégrés dans un processus d'inférence floue afin de reconnaître l'identité des objets détectés. Dans ce processus d'inférence floue, des connaissances qualitatives relatives aux objets et associées à ces notions d'intensités et de textures sont données par l'expert. Ces connaissances sont représentées par des fonctions d'appartenance trapézoïdales définies en utilisant des méthodes heuristiques. Dans cette méthode, on insiste sur le rôle des connaissances d'expert afin de reconnaître les objets dans la scène. En revanche, cette intervention croissante de l'expert, par l'introduction des concepts heuristiques pour fonctionner avec différents jeux de données, rend cette méthode très sensible à ce type de connaissances.

B. Gong *et al.* [45] ont proposé deux signatures sémantiques pour améliorer les méthodes de recherche d'objets 3D basées sur les primitives liées à la forme d'objet. La première signature sémantique consiste à émuler la description humaine des objets en termes d'attributs et la deuxième signature consiste à décrire un objet en le comparant avec les classes d'objets connues *a priori*. Ces deux signatures permettent de mettre en correspondance les primitives de forme de bas niveau avec des descriptions sémantiques de haut niveau relatives aux objets. Cette coopération entre les informations de bas niveau et les connaissances d'expert du haut niveau permet d'améliorer la reconnaissance des objets.

D'une manière générale, les méthodes d'interprétation proposées au niveau objet dépendent fortement de la coopération entre les informations extraites de bas niveau d'analyse de scène et les connaissances *a priori* liées à l'expert. Malgré toutes les tentatives d'automatisation du processus d'extraction des objets dans les scènes, le rôle de l'expert reste crucial afin de réaliser une reconnaissance fiable de ces objets. Encore une fois, nous constatons à ce niveau, l'importante capacité de la représentation floue pour l'intégration des connaissances d'expert dans le processus d'interprétation.

1.5.5 Le niveau scène

C'est le niveau sémantique le plus élevé d'un système de vision par ordinateur dont la tâche principale est l'identification de l'ensemble des objets existants dans la scène et la description globale de cette scène (structures, objets, et ses relations) « *au-delà de la reconnaissance d'objets isolés* » [44]. Ce niveau d'interprétation se caractérise par la nécessité de disposer d'un ensemble plus large de connaissances *a priori* principalement qualitatives portant d'une part, sur les différents types d'objets présents dans la scène et d'autre part, sur le contexte spatial / temporel de ces objets [45] [46]. En effet, de telles connaissances sont très importantes pour obtenir une description précise de la scène. Par exemple, une relation spatiale permet soit de valider la présence d'un objet par l'existence d'autres objets déjà détectés dans son contexte réel (un aéroport est généralement à côté d'une ville, une voiture est sur une route, etc.), soit d'inférer une nouvelle hypothèse d'objet (une ville peut conduire à l'existence d'un aéroport) [7]. L'exploitation de la diversité des connaissances *a priori* disponibles ainsi que le raisonnement mené influencent le choix de la représentation de ces connaissances. En effet, cette représentation doit être flexible pour permettre d'un côté, l'intégration autant que possible des connaissances de l'expert, et d'un autre côté, la manipulation des différents types d'imperfections intrinsèquement liées à ces connaissances [46]. Il existe plusieurs mécanismes de représentation des connaissances pour ce haut niveau tels que les ontologies [47], les règles de production [48], les réseaux sémantiques [49], etc. Le choix d'un mécanisme de représentation dépend fortement de l'objectif visé, des données disponibles et des contraintes liées à l'application considérée.

Plusieurs systèmes d'interprétation au niveau scène ont été proposés dans la littérature. Ils diffèrent par un certain nombre d'aspects tels que les stratégies d'interprétation, la représentation des connaissances ou le domaine d'application. Parmi ces travaux, on peut citer le système d'interprétation sémantique d'image proposé par C. Hudelot [1]. Ce système est basé sur la vision cognitive et se focalise sur la proposition de solutions génériques et indépendantes de toute application. L'architecture proposée est une architecture distribuée qui se base sur la coopération de trois systèmes à base de connaissances (SBCs). Pour chaque SBC, un modèle générique a été proposé en formalisant la connaissance et des stratégies de raisonnement dédiées. Deux ontologies ont été utilisées pour faciliter l'acquisition de la connaissance et permettre l'interopérabilité entre les trois différents SBCs. Les solutions proposées ont été validées sur une application de diagnostic précoce des pathologies végétales et en particulier des pathologies du rosier de serre. Deux inconvénients majeurs de ce système de vision cognitive sont : 1) les relations entre les concepts visuels associés à une classe et les descripteurs numériques d'images, sont

codées manuellement, et 2) ce système ne tient pas compte des imperfections liées aux connaissances disponibles. Afin de réduire le fossé entre les connaissances symboliques et les données numériques, une solution proposée en prolongement du travail de [1] consiste à recourir à des méthodes de reconnaissance de formes.

P. Garnesson et al [7] ont proposé un système multi-spécialistes pour l'interprétation de scène (MESSIE : *Multi Expert System for Scene Interpretation and Evaluation*). Ce système dispose d'une architecture du type tableau noir « *blackboard* » et est composé de trois niveaux : 1) niveau image qui contient les spécialistes d'extraction des primitives et le choix des algorithmes de traitement d'image, 2) niveau objet qui fournit un ensemble de spécialistes, chacun d'entre eux étant spécialisé dans la reconnaissance d'une classe d'objets, et 3) niveau scène qui permet de guider l'interprétation de la scène. Ce sont les spécialistes qui décident quels objets chercher et dans quelles zones de l'image. Ils peuvent être déclenchés à la demande d'autres spécialistes ou automatiquement lorsqu'un événement intéressant se produit sur le tableau noir. Les objets dans la scène ont été représentés d'une façon générique (indépendante de type de capteur en vue d'une application future dans le cadre de la fusion multi-capteurs). Cette représentation est basée sur quatre points de vue (géométrique, radiométrique, contexte spatial et fonctionnalité) et réalisée en utilisant des règles et des réseaux sémantiques. Chaque objet est lié à un spécialiste qui dispose de sa propre stratégie de reconnaissance. La détection des objets est mise en œuvre en deux étapes : 1) des objets saillants (objets macroscopiques) sont d'abord recherchés, puis 2) ces objets sont utilisés pour soutenir la détection de petits objets. Les relations spatiales entre les objets sont principalement prises en compte pour valider (ou rejeter) les hypothèses d'objets. Pour gérer les imperfections (l'incertitude et l'imprécision) des modèles et des données, MESSIE utilise la théorie des ensembles flous et la théorie des possibilités. Ce système a été validé sur une application d'interprétation de scène en imagerie aérienne.

Un point clé de ce système est la représentation des objets en se basant sur les quatre points de vue (géométrique, radiométrique, contexte spatial et fonctionnalité) et d'une façon indépendante du type de capteur. Un autre point clé réside dans l'utilisation des représentations floue et possibiliste afin de gérer les différents types d'imperfections liées aux connaissances disponibles. Un inconvénient majeur de ce système est l'utilisation intensive des connaissances d'expert, ce qui rend la conception de ce système très liée à l'application visée.

T. Matsuyama and V. Hwang [17] ont proposé un système d'interprétation de scène appelé SIGMA : *A knowledge-based aerial image understanding system*). Ce système repose sur une hiérarchie de trois systèmes experts de raisonnement : 1) l'expert de bas niveau de vision (LLVE : Low Level Vision Expert) qui fournit des connaissances liées aux méthodes d'analyse d'image, 2) l'expert de sélection de modèle (MSE : Model Selection Expert) qui possède les connaissances de correspondance entre les données image (bas niveau) et la scène (des objets), et 3) le système d'expert de raisonnement géométrique (GRE : Geometric Reasoning Expert) pour le raisonnement spatial entre les objets. Les différents systèmes experts communiquent entre eux par des requêtes. La représentation des connaissances est basée sur le concept de classe d'objets. Une classe d'objets représente les propriétés et structures d'un objet du domaine d'application ainsi que les connaissances liées à la façon de reconnaître cet objet (règles de raisonnement spatial). Afin de vérifier l'existence d'objets hypothétiques, LLVE est d'abord activé par le module de reconnaissance d'objets. Le raisonnement spatial entre ces objets est réalisé par le système GRE qui effectue une accumulation d'évidences en intégrant des processus descendant et ascendant permettant de tester les objets, d'évaluer leur consistance et de les regrouper. Ce système propose une architecture claire et bien formalisée pour l'interprétation de scène. En effet, cette architecture permet la séparation de trois types de connaissances 1) les connaissances liées à l'image, 2)

les connaissances liées à la scène et 3) les connaissances permettant la mise en correspondance entre la scène et l'image. Un autre point clé de ce système consiste en l'utilisation du processus descendant qui permet d'un côté (en coopération avec le processus ascendant) d'accumuler des évidences concernant l'existence de ces objets, et d'un autre côté, d'exploiter les connaissances d'expert liées à la segmentation LLVE. Ce processus descendant permet de réduire les imperfections liées à la segmentation. Un inconvénient majeur de ce système est l'acquisition des connaissances nécessaires pour la mise en œuvre de ce système.

Généralement, la majorité des systèmes d'interprétation au niveau scène présentés précédemment sont composés de plusieurs niveaux sémantiques (généralement trois niveaux) allant des bas niveaux vers les hauts niveaux. Les connaissances *a priori* et les connaissances d'expert sont primordiales aux différents niveaux. Nous constatons aussi l'importance de l'exploitation du contexte spatial des objets dans ces systèmes d'interprétation. Le lien entre ces niveaux constitue l'un des problèmes majeurs rencontrés lors de la conception de tels systèmes ainsi que l'acquisition des connaissances nécessaires pour le processus d'interprétation. L'adaptation des processus ascendant et descendant dans le système d'interprétation de scène (par exemple, le système SIGMA) semble être la plus prometteuse afin de remédier au problème du lien entre les différents niveaux sémantiques. Néanmoins, ces systèmes ne fournissent pas les mécanismes permettant à l'expert d'explorer plus finement (ou de remettre en cause), le contenu informationnel des objets détectés au plus haut niveau d'interprétation.

1.6 Conclusions

Dans ce chapitre, nous avons présenté, dans un premier lieu, le cadre général pour toute tâche d'interprétation de scène. Ensuite, nous avons décrit le processus d'interprétation de scène et les enjeux relatifs. Nous avons particulièrement souligné l'importance des connaissances *a priori* dans l'identification du contenu sémantique pour les différents niveaux d'interprétation de scène : sous-pixelique, pixelique, région, objet et scène. Nous avons également montré la nécessité de manipuler les divers types d'imperfections affectant le processus d'interprétation afin d'avoir une interprétation plus fiable et plus robuste de la scène. Un bref rappel de certaines méthodes de reconnaissance de formes a été présenté. Ces méthodes sont souvent appliquées aux trois premiers niveaux sémantiques d'interprétation de scène : sous pixelique, pixelique et région. En fonction de l'objectif final ou du besoin exprimé par l'expert, nous avons dressé un état de l'art autour des systèmes existants en les catégorisant selon les cinq niveaux sémantiques d'interprétation. Un regard particulier a été porté sur les niveaux d'interprétation de scène les plus bas (sous pixelique, pixelique, région) et les problèmes qui y sont associés car ces niveaux correspondent à l'objectif de ce travail de thèse. L'étude de ces systèmes nous permet de déduire les remarques suivantes :

- Les systèmes d'interprétation de scène concernant les "bas niveaux" opèrent au niveau pixelique soit pour aboutir au niveau sous-pixelique (démixage, par exemple), soit pour aboutir au niveau région, sans formaliser les liens entre ces trois niveaux. Pourtant, un tel lien ne peut qu'améliorer les résultats globaux de l'interprétation.
- L'importance du rôle de l'expert a été montrée dans la majorité de ces systèmes. Néanmoins, le lien entre les connaissances de l'expert et les mesures numériques permettant d'extraire le contenu sémantique de la scène constitue un problème récurrent lors de la conception d'un système d'interprétation de scène. Notons à ce stade, la difficulté pour la représentation et l'intégration des connaissances d'expert dans le contexte où les classes thématiques contenues dans la scène sont très difficiles à discriminer (image très bruitée ou fortement texturée, par exemple).

- Une tendance importante dans la majorité des systèmes d'interprétation consiste à utiliser plusieurs sources de connaissances (plusieurs capteurs, des sources de connaissances additionnelles, ...) et ceci afin d'améliorer les résultats d'interprétation et les rendre plus fiables et plus robustes.
- Un intérêt croissant pour l'utilisation de la théorie des ensembles flous et la théorie des possibilités dans le processus d'interprétation a été constaté. Cet intérêt est motivé par la flexibilité de ces deux théories pour la représentation et l'intégration des connaissances disponibles et pour la fusion des connaissances de nature hétérogène et entachées d'ambiguïtés et d'incertitudes.
- La majorité des systèmes d'interprétation proposés reposent sur des connaissances *a priori* qui sont supposées représentatives du contenu informationnel à identifier (connaissances qualitatives, échantillons d'apprentissage, données multi-capteurs). L'acquisition de telles connaissances lors du processus d'interprétation constitue l'un des problèmes majeurs rencontrés pour la conception des systèmes d'interprétation de scène. Ceci montre l'importance de doter le système d'interprétation d'un mécanisme d'extraction de connaissances à partir des données elles-mêmes avec une évaluation de la certitude des connaissances extraites.

En se basant sur les remarques précédentes concernant les systèmes existants d'interprétation de scène, nous proposons une approche d'interprétation de scène qui permet de tirer profit des avantages de certains systèmes (représentation et diffusion itérative des connaissances dans l'espace possibiliste [111]) et de remédier aux inconvénients d'autres systèmes, surtout les systèmes opérant aux niveaux d'interprétation sous pixelique, pixelique et région (manque d'une démarche permettant de remettre en cause, par l'expert, du contenu informationnel des régions identifiées au niveau région). En effet, ce travail de recherche s'inscrit dans le contexte de la représentation des connaissances en vue de concevoir un système d'interprétation de scènes dans un environnement multi-sources de connaissances. Etant donné les différents types d'imperfections liées aux sources de connaissances, nous nous focalisons sur une représentation possibiliste des connaissances afin d'exploiter le maximum d'informations disponibles, en termes de caractéristiques et de structure spatiale. En effet, ce formalisme permet d'intégrer et de manipuler explicitement et au sein du même cadre, les différentes formes d'imperfection d'informations (liées aux capteurs, aux conditions d'acquisitions, aux observations et connaissances d'experts, etc.) ainsi que des connaissances *a priori* portant sur la scène analysée.

Dans ce travail, nous avons aussi adopté une analyse hiérarchique/sémantique de scène où nous étudions l'application de la théorie des possibilités dans les différents niveaux sémantiques (sous-pixelique, pixelique, région, objet et scène). Cette analyse hiérarchique repose d'une part sur un processus ascendant permettant d'accumuler l'évidence sur l'existence des régions (ou objets), et d'autre part, sur un processus descendant permettant de mettre en cause par l'expert, le contenu informationnel des objets détectés précédemment.

L'objectif final de ce travail de thèse consiste à mettre en place les outils possibilistes nécessaires à l'exploitation conjointe de plusieurs sources de connaissances afin d'améliorer l'interprétation des données acquises en vue d'une représentation plus riche de la scène observée : les deux principales applications visées sont donc la segmentation et le démixage. Ainsi, dans le processus de segmentation, un aspect important du raisonnement humain a été modélisé, en intégrant de manière itérative, l'information issue du contexte spatial des pixels. L'idée sous-jacente est de construire une base d'apprentissage représentative (par raffinement) des différentes classes d'une part, et de réduire l'incertitude liée à la classification des pixels lors de la diffusion des connaissances possibilistes d'autre part. La coopération entre le raffinement de classes et la diffusion des connaissances possibilistes permet de

discriminer les classes dans un contexte où ces classes sont difficiles à discriminer (faible contraste, environnement très bruité ou existence de classes multimodales). Dans l'étape de démixage, le contexte spatial des pixels est exploité afin d'identifier le contenu informationnel des régions détectées dans le processus de segmentation. Afin de détailler l'approche proposée, il est nécessaire de caractériser les imperfections et les méthodes permettant de les modéliser. Ceci sera détaillé dans le chapitre 2.

Chapitre 2

IMPERFECTION DE L'INFORMATION - CADRES MATHEMATIQUES

2.1 Introduction

L'analyse d'une scène constitue une phase essentielle dans le processus d'interprétation de cette scène. Cette phase consiste en l'extraction d'informations contenues dans les données issues potentiellement de plusieurs sources de connaissances. Ces informations sont pour la plupart imparfaites, c'est-à-dire qu'elles comportent des incertitudes, imprécisions, incomplétudes, ambiguïtés, voire qu'elles puissent être erronées. Il faut noter que ces imperfections accompagnent également toutes les étapes du processus d'extraction d'informations à partir des sources de connaissances, depuis les phénomènes observés jusqu'aux traitements et d'interprétation [46]. Par conséquent, pour qu'une interprétation de scène soit cohérente, riche et fiable, il faut prendre en considération ces imperfections dans les différentes phases de traitement et d'analyse.

Dans ce chapitre, nous abordons, dans un premier temps, la définition et la caractérisation de l'information ainsi que les différentes formes d'imperfections pouvant l'entacher. Ensuite, nous présentons les aspects essentiels des différentes approches utilisées pour représenter et traiter les informations imparfaites. La théorie des possibilités est particulièrement détaillée car elle sera considérée comme le cadre général du système d'interprétation de scène que nous proposons. En fin de chapitre, nous discuterons de l'intérêt fondamental que représente la fusion d'informations imparfaites en vue d'obtenir une information finale plus fiable.

2.2 Caractérisation d'information (imperfection et hétérogénéité)

En traitement de l'information, les trois termes essentiels *donnée*, *information* et *connaissance* sont souvent utilisés pour désigner des concepts qui se chevauchent. En effet, ces termes ne sont pas définis

précisément et la frontière entre ces termes est floue et souvent subjective. Plusieurs définitions de ces termes ont été proposées dans la littérature [50]. La principale différence entre ces termes réside dans le niveau d'abstraction considéré. Les *données* (normalement brutes) sont considérées comme le niveau le plus bas d'abstraction, l'*information* (donnée traitée et portant un sens informatif) est le niveau suivant et enfin, la *connaissance* (information synthétisée) constitue le niveau le plus élevé parmi ces trois niveaux [51].

Les données brutes ne portent pas de signification particulière. Pour que ces données deviennent informatives, elles doivent prendre un sens dans le contexte de leur utilisation. Par exemple, la valeur « $x = 39$ » ne représente qu'un entier positif d'un point de vue mathématique purement abstrait. Cette valeur peut être transformée afin qu'elle devienne plus informative en lui associant un sens physique tel que « $x = 39^\circ\text{C}$ », x est alors considérée comme un élément dans un ensemble ayant un sens sémantique. Finalement, en précisant le contexte lié à « $x = 39^\circ\text{C}$ » (température ambiante ou température d'un patient, par exemple), cette valeur devient alors plus informative et nous pouvons même la considérer comme étant « une information ».

2.2.1 Définition d'une information

Une information est définie (Figure 2.1) [51] comme un triplé (Θ, Ω, Inf) comportant deux ensembles Θ, Ω et une fonction Inf associant ces deux ensembles.

L'ensemble Θ , appelé *ensemble de définition de l'information*, est formé des objets sur lesquels porte cette information (ensemble de patients, ensemble de zones ayant les dimensions des cellules de résolution d'un capteur donné, etc.).

L'ensemble Ω , appelé *ensemble de contenus informationnels*, est formé des valeurs informatives délivrées par l'information (températures, labels, tensions électrique, décisions, valeurs de réflectance mesurées par un capteur, etc.).

La fonction $Inf(.)$ est appelée *fonction informative*. Elle représente la mesure d'une quantité physique ou simplement un modèle mathématique associant un objet de l'ensemble Θ à un ou plusieurs contenus informationnels de l'ensemble Ω . Chaque correspondance entre un objet $\Theta_m \in \Theta$ et un contenu informationnel $x_n \in \Omega$ réalisée via la fonction $Inf(.)$ est appelée un *élément d'information*.

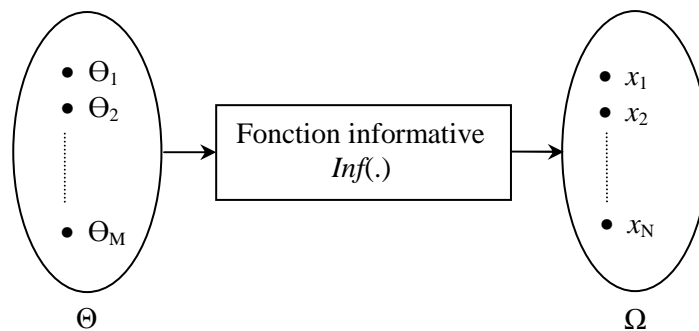


Figure 2.1 : Définition et structure de l'information

Considérons, à titre d'exemple, le cas d'une image numérique. La formation d'une telle image implique l'utilisation d'un capteur réalisant la mesure d'une quantité physique (réflectance, absorption, radiance, etc.) sur une scène du monde réel. Ce capteur opère sur des objets élémentaires, appelés *cellules de résolution*, ayant des dimensions « imposées » par les contraintes de ce capteur. Ces objets élémentaires sont représentés sous la forme de pixels dans l'image en sortie du capteur. Le résultat de la mesure, réalisée par le capteur au niveau de chaque cellule de résolution, est présenté sous la forme d'un

niveau de gris associé avec chaque pixel de l'image numérique. Par conséquent, l'ensemble de définition associé à cette information d'imagerie est simplement constitué par l'ensemble des cellules de résolution (ayant les dimensions imposées par le capteur). L'ensemble des contenus informationnels est formé par les niveaux de gris de l'image, et finalement, la fonction informative est constituée par le processus de mesure de la quantité physique pour laquelle ce capteur est destiné. A titre d'exemple, la Figure 2.2 illustre un exemple d'une telle information en télédétection (capteur multi-spectral SPOT en mode panchromatique) correspondant à une image optique de l'île de la Réunion.

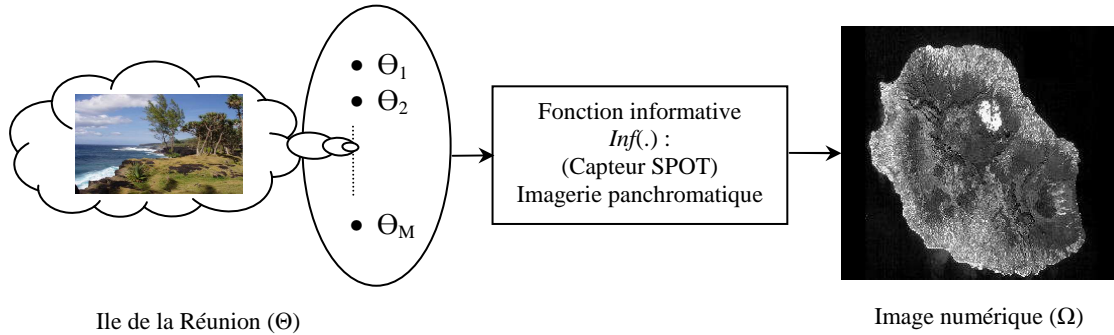


Figure 2.2 : Exemple d'une information en télédétection

Il est important de noter que la connaissance de la fonction informative est cruciale pour réaliser les tâches de haut niveau de traitement d'information, aussi appelées *tâches intelligentes*, comme l'interprétation de scène, la fusion d'informations, la fouille de connaissances, etc.

A titre d'exemple, lorsqu'un capteur actif est utilisé (capteur radar en télédétection, par exemple), l'interprétation, en termes de contenu sémantique, d'une zone sombre dans l'image observée, correspond à une zone de non-réflectance (une zone lisse, c'est-à-dire que la longueur d'onde utilisée est largement supérieure aux variations de la surface terrestre illuminée), ou simplement à une zone d'ombre. En revanche si le capteur utilisé est un capteur passif (multi ou hyper spectral), alors l'interprétation sémantique d'une zone sombre est complètement différente. Cette interprétation revient à chercher les classes thématiques qui absorbent l'onde électromagnétique incidente dans la bande spectrale concernée.

2.2.2 Caractérisation d'une information

Du point de vue de la modélisation mathématique, une information peut être caractérisée par trois propriétés principales : l'exhaustivité, l'exclusivité et la force du lien [74].

Une information est dite *exhaustive* si $\forall \Theta_m \in \Theta$, il existe un élément $x_n \in \Omega$ qui vérifie le lien $Inf(\Theta_m) = x_n$. En d'autres termes, l'exhaustivité signifie que l'ensemble Ω contient tous les contenus informationnels possibles et qui sont produits par la fonction informative Inf .

Une information est dite *exclusive* si $\forall \Theta_m \in \Theta$, il existe un et seulement un, contenu informationnel $x_n \in \Omega$ qui vérifie l'association $Inf(\Theta_m) = x_n$. Ceci signifie qu'un objet ne peut pas être associé à deux contenus informationnels à la fois.

La troisième propriété, la force du lien, se focalise sur la véracité de l'association entre les objets de l'ensemble de définition et les contenus informationnels. En effet, une information peut avoir soit un lien d'association binaire dans lequel $Inf(\Theta_m) = x_n$ est une association totalement vraie ou totalement fausse, soit un lien ayant une véracité partielle pour laquelle $Inf(\Theta_m) = x_n$ est associée à un degré de vérité partielle. Cette notion de lien partiel constitue la caractéristique fondamentale permettant de « sortir » des cloisons imposées par les théories mathématiques qui n'opèrent que dans un monde binaire de véracité.

En se basant sur ces trois propriétés, deux types d'information peuvent être distingués : une information *parfaite* et une information *imparfaite*.

En effet, une information est dite *parfaite* si et seulement si elle est exhaustive, exclusive et sa fonction informative réalise un lien binaire. En d'autres termes, le contenu informationnel $x_n \in \Omega$ associé à l'objet $\Theta_m \in \Theta$, est connu avec une totale certitude.

Toute information qui ne vérifie pas ces conditions est dite *imparfaite*.

2.2.3 Notion d'information imparfaite

La définition que nous venons de donner d'une information imparfaite est certainement très générale et reflète le nombre impressionnant de termes utilisés dans la littérature pour décrire les différentes formes d'imperfection. Néanmoins, la définition générale donnée d'une information nous permet de distinguer deux dimensions pour décrire l'imperfection d'une information. La première dimension concerne *l'imperfection intrinsèque* liée à la structure de l'information elle-même, et la deuxième dimension concerne *l'imperfection extrinsèque* liée au contexte d'utilisation de l'information.

Avant de décrire et de modéliser d'une façon précise l'imperfection intrinsèque, nous allons montrer quelques exemples de formes d'imperfections extrinsèques car cette dimension est fortement liée au contexte et à l'objectif de l'utilisation de l'information.

L'*incomplétude* est l'une des formes d'imperfection extrinsèque d'une information. Il s'agit du fait qu'une information, même parfaite, n'est pas « assez informative » pour un processus décisionnel dans lequel cette information est fournie par une source de connaissance. Ce manque « d'informativité » rend cette information incomplète pour atteindre l'objectif visé. Dans une telle situation, les informations apportées par d'autres sources de connaissances sont incontournables afin d'augmenter la fiabilité du processus décisionnel global. A titre d'exemple, en imagerie optique, les objets se trouvant dans une zone d'ombre ne peuvent pas être détectés et l'utilisation d'autres sources d'information est incontournable.

Une autre forme d'imperfection extrinsèque d'une information est celle du *conflit*. Cette forme d'imperfection est rencontrée lorsque plusieurs informations sont exploitées dans un processus de prise de décision. Deux informations sont dites en conflit si elles portent des contenus informationnels contradictoires pour la prise d'une décision [53]. A titre d'exemple, une information décisionnelle attribuant avec certitude, une classe à un objet est en « total » conflit avec une autre information décisionnelle attribuant avec certitude, une classe différente à ce même objet. Notons à ce stade, l'utilisation d'une échelle qualifiant le degré de conflit entre deux informations allant du totalement conflictuel vers le total accord.

Concernant l'imperfection intrinsèque d'une information, il existe deux modèles de base de ces imperfections : *l'incertitude* et *l'ambiguïté*. Nous décrivons brièvement ci-dessous ces deux types d'imperfections.

2.2.3.1 L'incertitude

En considérant une information qui vérifie les trois propriétés de base : l'*exhaustivité*, l'*exclusivité* et le *lien total* (l'association entre un contenu informationnel $x_n \in \Omega$ et un objet $\Theta_m \in \Theta$, est soit totalement vraie ou totalement fausse).

Etant donné un objet $\Theta_m \in \Theta$, cette information est dite *incertaine* si le contenu informationnel $x_n \in \Omega$ qui est associé avec l'objet $\Theta_m \in \Theta$ est inconnu. En d'autres termes, l'incertitude est relative à la vérité « binaire » de l'association entre un objet de l'ensemble de définition de cette information et le contenu informationnel qui est supposé être en lien avec cet objet. Le résultat d'un tirage aléatoire, la décision à prendre face à une situation et la classe d'un objet dans un problème de reconnaissance de formes

constituent des exemples de cas pratiques pour lesquels on est confronté à la forme d'incertitude de l'imperfection intrinsèque d'une information.

Les différentes approches permettant le traitement d'informations incertaines (théorie des probabilités, théorie des croyances et la théorie des possibilités) se différencient par leur mode de représentation des connaissances disponibles et de description de l'incertitude, ainsi que par les outils de traitement et de raisonnement permettant de manipuler des informations incertaines afin de « lever l'incertitude » et arriver ainsi à une décision certaine d'association entre un objet $\Theta_m \in \Theta$ et l'unique contenu informationnel $x_n \in \Omega$ qui lui est associé.

Notons à ce stade que l'imprécision n'est qu'un cas particulier de l'incertitude. En effet, une information est dite imprécise lorsqu'elle est affectée par l'incertitude et lorsque les connaissances disponibles concernant le contenu informationnel recherché, $x_n \in \Omega$, sont données sous la forme d'un sous-ensemble $A \subseteq \Omega$ de contenus informationnels.

L'*imprécision* réfère donc à un défaut quantitatif de connaissance qui ne permet pas de décrire précisément l'unique contenu informationnel. Cette forme d'imperfection est souvent rencontrée en intégrant les erreurs dans les systèmes de mesures physiques. Néanmoins, l'imprécision est aussi rencontrée dans les systèmes décisionnels. Par exemple, étant donné un système de reconnaissance de formes où $\Omega = \{C_1, C_2, C_3, C_4, C_5\}$ désigne l'ensemble composé de cinq classes qui sont supposées mutuellement exhaustives et exclusives, alors être dans l'incapacité d'associer un objet à l'une des deux classes C_1 ou C_3 représente une information imprécise parce que la vraie classe de l'objet considéré n'est connue qu'au travers du sous-ensemble $\{C_1, C_3\} \subseteq \Omega$.

2.2.3.2 L'ambiguïté

La majorité des travaux de recherche s'est focalisée sur l'incertitude. Pourtant, plusieurs situations réelles ne vérifient pas les conditions de base de l'incertitude et, malgré ce « défaut », les outils adaptés au traitement de l'incertitude ont été appliqués d'une façon abusive à l'autre forme d'imperfection de l'information. Considérons, par exemple, la classification pixelique dans le cas de pixels comportant plusieurs contenus thématiques. Dans ce cas, nous sommes confrontés à la « contradiction conceptuelle » d'affecter ces pixels à une seule classe parmi les classes thématiques (ce qui ne correspond pas à la réalité).

Les différentes solutions proposées pour traiter ce type de situations en appliquant les approches classiques de traitement de l'incertitude (essentiellement les approches probabilistes) ne sont que des solutions artificielles qui cherchent à contourner la deuxième forme d'imperfection intrinsèque de l'information qui est l'ambiguïté.

Cette forme d'imperfection concerne l'information *exhaustive*, mais pas nécessairement *exclusive* et pouvant avoir des *liens partiels*, c'est-à-dire que chaque objet $\Theta_m \in \Theta$ peut être associé simultanément à plusieurs contenus informationnels de Ω , avec des degrés non forcément binaires de vérité.

La théorie des ensembles flous a été proposée afin d'apporter un outil permettant la représentation des connaissances ainsi que des mécanismes de raisonnement afin de traiter cette forme d'imperfection.

En effet, et étant donné qu'une telle information n'est pas exclusive et que le lien entre un objet et un contenu informationnel n'est pas forcément binaire, la théorie des ensembles flous considère chaque objet de l'ensemble de définition comme étant en lien avec « tous » les contenus informationnels avec une force de lien variant entre « 0 » et « 1 ». Cette véracité ou force, de lien partielle devient une fonction définie sur l'ensemble des contenus informationnels et elle est appelée la *fonction d'appartenance*.

Notons qu'au-delà des différentes formes d'imperfections rencontrées, les contenus informationnels peuvent être de type quantitatif (numérique ou binaire), qualitatif (symbolique et ordinal), voire des contenus hybrides (une information symbolique codée par des nombres) [54]. Ainsi, cette hétérogénéité d'informations doit être prise en compte lors du traitement d'informations.

2.3 Approches de modélisation d'informations imparfaites

La modélisation concerne le passage d'informations extraites des sources de connaissances vers une représentation mathématique liée à une approche ou une théorie, particulière. Cette modélisation est une tâche indispensable dans les différentes étapes d'un système d'interprétation de scène. Une bonne modélisation permet à un tel système de gérer explicitement les différentes formes d'imperfections rencontrées et de combiner les informations afin d'obtenir des informations plus fiables et améliorer ainsi leurs interprétations.

Différentes théories et approches ont été proposées dans la littérature pour modéliser les imperfections. Les principales approches sont : la théorie des probabilités, la théorie des fonctions de croyance, la théorie des ensembles flous et la théorie des possibilités. Chacune de ces théories concerne un type d'imperfection et un contexte particulier d'application.

Dans ce paragraphe, nous présentons brièvement ces théories tout en mettant l'accent sur le type d'imperfection modélisée. Un intérêt particulier sera porté sur la théorie des possibilités car elle constitue le cadre que nous avons adopté dans notre travail de recherche.

Nous allons noter dans la suite de ce paragraphe, l'ensemble $\Omega = \{x_1, x_2, \dots, x_N\}$ qui représente un univers de contenus informationnels comportant N éléments que l'on appelle *singletons* (valeurs observables, classes, hypothèses, décisions élémentaires, etc.).

2.3.1 Théorie des probabilités

La théorie des probabilités est la théorie la plus répandue pour modéliser les informations affectées par l'incertitude [55] (c'est-à-dire un et un seul singleton peut se produire), l'objectif étant de « découvrir » cet unique singleton. Dans le cadre de cette théorie, la connaissance liée à la réalisation des singletons x_n est représentée par une distribution de probabilités $\text{Pr}(\cdot)$. Cette distribution associe à chaque singleton x_n de Ω , une valeur $\text{Pr}(x_n)$ désignant la probabilité pour que le singleton x_n soit l'unique singleton qui s'est produit [56] :

$$\begin{aligned} \text{Pr}: \Omega &\rightarrow [0,1] \\ x_n &\rightarrow \text{Pr}(x_n) \end{aligned} \tag{2.1}$$

L'occurrence d'un événement $A \subseteq \Omega$ (A est un sous-ensemble de Ω) signifie que l'unique singleton qui s'est produit appartient à cet événement. Cette occurrence est caractérisée par une mesure $\text{Pr}(A)$ dite *mesure de probabilité*, qui est la somme des probabilités élémentaires affectées aux singletons formant cet événement. En d'autres termes, la valeur de cette mesure $\text{Pr}(A)$ peut être interprétée comme le degré de croyance pour que l'unique singleton qui s'est produit, appartienne au sous-ensemble A . Cette mesure de probabilité $\text{Pr}(A)$ est définie comme une mesure ensembliste qui vérifie les propriétés suivantes :

- $\text{Pr}(\emptyset)=0$ et $\text{Pr}(\Omega) = \sum_{x_n \in \Omega} \text{Pr}(x_n) = 1$;
- $\forall A, B \subseteq \Omega, \text{Pr}(A \cup B) = \text{Pr}(A) + \text{Pr}(B) - \text{Pr}(A \cap B)$;
- $\text{Pr}(A) + \text{Pr}(A^C) = 1$, avec A^C l'événement complémentaire de A .

La théorie des probabilités repose sur des bases mathématiques solides et riches et elle constitue un outil puissant pour représenter fidèlement les informations entachées par l'incertitude. Néanmoins, cette théorie ne permet pas la modélisation du raisonnement humain où les informations et les connaissances

fournies sont de nature ambiguë [57]. A titre d'exemple, la réalisation de deux singletons à la fois (comme c'est le cas des pixels comportant plusieurs classes thématiques, i.e. des mixels), constitue un cas d'ambiguïté pour lequel la théorie des probabilités propose des solutions artificielles.

2.3.2 Théorie des fonctions de croyance

La théorie des fonctions de croyance, aussi appelée théorie de Dempster-Shafer, est issue des travaux de Dempster en 1967 sur les probabilités inférieures et supérieures [58], et ensuite reprise par Shafer sous le nom de la *théorie de l'évidence* [59]. Elle est devenue un outil très populaire en intelligence artificielle pour aborder les problèmes liés à la représentation des connaissances et à la prise de décision. Cette théorie traite le cas où l'univers des contenus informationnels Ω est exhaustif et exclusif et permet de modéliser à la fois l'incertitude et l'imprécision exprimée d'une façon ensembliste. La modélisation de cette forme d'imperfection se réalise au travers de deux fonctions, l'une dite de *crédibilité* $Cr(.)$ et l'autre dite de *plausibilité* $Pl(.)$, toutes deux dérivées d'une fonction de masses de croyance $m(.)$.

Dans cette théorie, les fonctions de masses de croyance sont définies sur tous les sous-ensembles A (événements) possibles de l'univers Ω et pas seulement sur les singletons comme dans la théorie des probabilités. La fonction $m(.)$ de masses de croyance est définie comme suit :

$$\begin{aligned} m: 2^\Omega &\rightarrow [0,1] \\ A &\rightarrow m(A) \end{aligned} \quad (2.2)$$

où $m(A)$ désigne la croyance en la réalisation de l'évènement A : c'est donc la croyance que le singleton unique qui s'est produit, appartienne effectivement à cet évènement A . Par ailleurs, cette fonction $m(.)$ vérifie les propriétés suivantes :

- $m(\emptyset)=0$;
- $\sum_{A \subseteq \Omega} m(A)=1$.

Il est important de noter que la masse de croyance est affectée à chaque sous-ensemble de singletons A d'une façon globale, et qu'il n'est pas possible de la répartir sur les singletons composant l'évènement A .

Ainsi, à partir de la fonction $m(.)$, on définit respectivement les fonctions de crédibilité $Cr(.)$ et de plausibilité $Pl(.)$ comme suit :

$$Cr(A) = \sum_{B \subseteq A} m(B) \quad (2.3)$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) = 1 - Cr(A^c) \quad (2.4)$$

où A^c représente l'évènement complémentaire de A . La fonction de crédibilité $Cr(A)$ peut être interprétée comme la croyance minimale en la réalisation de A tandis que la fonction de plausibilité $Pl(A)$ mesure la croyance avec laquelle on ne doute pas de la réalisation de A et peut être interprétée comme la croyance maximale en la réalisation de A . En effet, cette capacité de la théorie de Dempster-Shafer pour représenter le doute entre la réalisation des singletons permet de traiter des informations incertaines de nature probabiliste ainsi que certains types de l'imprécision, mais avec une interprétation différente de la théorie des probabilités (i.e. ensembliste). Néanmoins, cette théorie comme la théorie des probabilités, ne permet pas de traiter les informations ambiguës.

2.3.3 Théorie des ensembles flous

La théorie des ensembles flous, proposée par Zadeh en 1965 [60], constitue un outil mathématique puissant pour la modélisation, la description et la prise de décision dans le cas d'informations ambiguës, approximatives, imprécises ou vagues. Ces types d'informations imparfaites ont des limites mal définies et sont plus fréquentes dans le raisonnement humain. La théorie des ensembles flous tente donc d'imiter la façon avec laquelle l'humain manipule ces informations. Elle est basée sur l'idée d'appartenance partielle où plusieurs contenus informationnels peuvent se produire simultanément, mais aussi partiellement.

Nous allons faire un bref rappel des principaux concepts de cette théorie afin de montrer sa capacité de modélisation et de description des informations et concepts manipulés par l'humain. Rappelons que ces informations et concepts sont essentiels dans le processus d'interprétation de scène.

2.3.3.1 Définitions

Dans la théorie des ensembles classiques, l'appartenance d'un élément à un ensemble classique ou algébrique peut avoir deux états logiques : appartenir ou ne pas appartenir à cet ensemble. Zadeh a tenté de sortir de cette logique binaire en introduisant la notion d'appartenance partielle. Cette notion permet d'avoir des graduations dans l'appartenance d'un élément à un ensemble. Il s'agit d'autoriser un élément à appartenir plus moins fortement à un ensemble.

Soient Ω un univers exhaustif de contenus informationnels (i.e. de singletons) et A , une propriété caractérisant les singletons de Ω . En considérant le raisonnement humain, ces singletons peuvent vérifier cette propriété d'une façon partielle. Une telle propriété est appelée *vague* ou *ambiguë*. Cette propriété est représentée comme un ensemble flou A défini sur cet univers. Cet ensemble est caractérisé par une fonction d'appartenance μ_A définie par :

$$\begin{aligned} \mu_A : \Omega &\rightarrow [0,1] \\ x_n &\rightarrow \mu_A(x_n) \end{aligned} \quad (2.5)$$

où $\mu_A(x_n)$ désigne le degré d'appartenance du singleton x_n à l'ensemble flou A . Dans le cas où le nombre de singletons est raisonnable, on peut écrire cet ensemble flou comme suit :

$$A = \{(x_n, \mu_A) / x_n \in \Omega, \mu_A(x_n) \in [0,1]\} \quad (2.6)$$

Ainsi, un ensemble flou A est complètement défini sur l'univers Ω par sa fonction d'appartenance μ_A associant chaque singleton x_n de Ω à une valeur $\mu_A(x_n) \in [0,1]$ qui représente le degré d'appartenance du singleton x_n à l'ensemble A , c'est-à-dire le degré de vérité que x_n appartienne à A (ou vérifie cette propriété ambiguë).

Trois cas sont à distinguer :

- $\mu_A(x_n)=0$: x_n n'appartient pas à A ;
- $0 < \mu_A(x_n) < 1$: x_n appartient partiellement à A , et
- $\mu_A(x_n) = 1$: x_n appartient entièrement à A .

Bien que la fonction d'appartenance puisse prendre n'importe quelle forme (tant que cette fonction permet d'associer à chaque singleton de Ω une valeur de l'intervalle $[0,1]$), cinq types standards de fonction d'appartenance sont les plus fréquemment utilisés dans le cadre de cette théorie, il s'agit des fonctions S, fonctions L, fonctions triangulaires, fonctions trapézoïdales et fonctions gaussiennes [62]. La Figure 2.3 illustre les formes géométriques de ces cinq types de fonctions d'appartenance.

Notons que le choix de l'une de ces formes de fonctions d'appartenance dépend essentiellement des connaissances exprimées par l'expert. Une fois la forme de la fonction d'appartenance sélectionnée par l'expert, les paramètres associés à cette fonction sont déterminés soit manuellement par l'expert d'une façon empirique et directe, soit par l'utilisation d'un processus automatique d'optimisation appliqué sur des mesures floues (considérées comme des fonctions objectives) telles que l'entropie floue ou l'indice flou [84] [85].

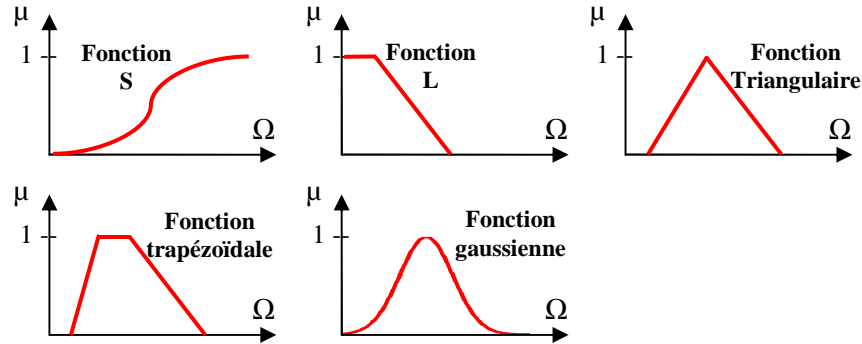


Figure 2.3 : Formes classiques de fonctions d'appartenance

Dans le cas où $\Omega = \{x_1, x_2, \dots, x_N\}$ désigne l'ensemble des niveaux de gris que l'on observe dans une image numérique, les formes classiques de fonctions d'appartenance sont utilisées pour la représentation des connaissances ambiguës liées à ces niveaux de gris, et exprimées souvent par un expert. La fonction S (resp. la fonction L) est couramment utilisée pour représenter le fort degré de « luminosité » ou de brillance (resp. un degré d'« obscurité ») des pixels. Les fonctions triangulaires, trapézoïdales et gaussiennes sont également utilisées pour décrire des plages de niveaux de gris dites « moyennes ». Considérons, à titre d'exemple, un système de classification d'images reposant sur un ensemble de règles floues [63]. Une règle potentielle de ce système peut être la suivante : « si le niveau de gris du pixel observé est *moyen* ALORS ce pixel doit être attribué la classe "végétaux" ». Si une fonction d'appartenance trapézoïdale de la Figure 2.4 est utilisée pour modéliser l'information ambiguë "niveau de gris *moyen*", les niveaux de gris entre 115 et 135 vérifient totalement cette règle en ayant des valeurs d'appartenance totale. Les niveaux de gris dans les intervalles [100, 115] et [135, 150] vérifient partiellement cette règle, et les niveaux de gris à l'extérieur de l'intervalle [100,150] ont une valeur d'appartenance nulle.

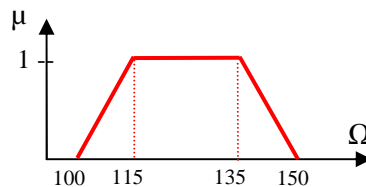


Figure 2.4 : Exemple de fonction d'appartenance trapézoïdale pour modéliser le concept "moyen"

2.3.3.2 Opérations sur les ensembles flous

Les opérations permettant de « combiner » les ensembles algébriques peuvent être généralisées aux ensembles flous. Considérons deux ensembles flous A et B de Ω caractérisés par deux fonctions d'appartenance μ_A et μ_B . Parmi les opérations définies sur les ensembles flous, on peut citer :

- **Complément** : le complément de l'ensemble A , noté A^C , est un ensemble flou défini par le biais de la fonction d'appartenance donnée pour tout x_n de Ω par:

$$\mu_{A^c}(x_n) = 1 - \mu_A(x_n) \quad (2.7)$$

- *Egalité* : les deux ensembles A et B sont dits égaux si et seulement si :

$$\forall x_n \in \Omega, \mu_A(x_n) = \mu_B(x_n) \quad (2.8)$$

- *Inclusion* : l'ensemble A est dit inclus dans B si et seulement si :

$$\forall x_n \in \Omega, \mu_A(x_n) \leq \mu_B(x_n) \quad (2.9)$$

- *Intersection* : l'intersection de deux ensembles flous A et B de Ω est un ensemble flou constitué des éléments de Ω affectés du plus petit des degrés d'appartenance à A et B . La fonction d'appartenance de $A \cap B$ est donc donnée pour tout x_n de Ω par :

$$\forall x_n \in \Omega, \mu_{A \cap B}(x_n) = \min(\mu_A(x_n), \mu_B(x_n)) \quad (2.10)$$

L'ensemble flou obtenu par l'opérateur *Intersection*, représente les éléments de Ω vérifiant les deux propriétés A et B simultanément, comme c'est illustré dans la Figure 2.5.

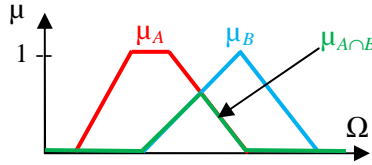


Figure 2.5 : Fonction d'appartenance de l'intersection de deux ensembles flous A et B

- *Union* : l'union de deux ensembles flous A et B de Ω est un ensemble flou constitué des éléments de Ω affectés du plus grand des degrés d'appartenance aux deux ensembles flous A et B (Figure 2.6). La fonction d'appartenance de $A \cup B$ est donnée pour tout x_n de Ω par

$$\forall x_n \in \Omega, \mu_{A \cup B}(x_n) = \max(\mu_A(x_n), \mu_B(x_n)) \quad (2.11)$$

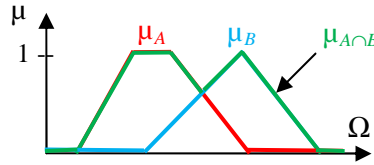


Figure 2.6 : Fonction d'appartenance de l'union de deux ensembles flous A et B

L'ensemble flou obtenu par l'opérateur *Union* représente les éléments de Ω vérifiant les propriétés A ou B ou les deux à la fois.

Après avoir donné un aperçu rapide sur quelques aspects essentiels de la théorie des ensembles flous et permettant de traiter les informations ambiguës, nous allons accorder une attention particulière à la théorie des possibilités. En effet, cette théorie permet de traiter et de fusionner les informations dans un contexte d'incertitude où les connaissances disponibles pour décrire cette incertitude sont de nature ambiguë et non probabiliste. Notons que ce type d'imperfection est très fréquemment présent dans les concepts et les informations manipulés par l'homme comme dans le cas sur lequel nous nous focalisons : l'interprétation de scène.

2.3.4 Théorie des possibilités

La théorie des possibilités est une théorie relativement nouvelle, destinée à la gestion de l'incertitude épistémique : c'est-à-dire, l'incertitude dans un contexte où les connaissances disponibles sont exprimées d'une façon ambiguë. A titre d'exemple, en cherchant à identifier la classe d'un pixel donné (contexte d'incertitude : ensemble exhaustif et exclusif des classes et la classe du pixel considéré n'est pas connue), une connaissance épistémique peut être celle exprimée par un expert en « décrivant le niveau de gris » du pixel avec le terme ambigu *sombre*.

Cette théorie a été introduite par Zadeh en 1978 comme une extension de la théorie des ensembles flous et de la logique floue [52]. Elle a ensuite été développée et décrite en profondeur par Dubois et Prade [69].

Dans la théorie des possibilités, l'incertitude liée à la réalisation d'un événement $A \subseteq \Omega$ (A peut être un sous-ensemble algébrique de Ω , ou même un ensemble flou défini sur Ω) est exprimée par deux mesures ensemblistes : la mesure de possibilité et celle de nécessité [70]. L'utilisation de ces deux mesures dans la théorie des possibilités permet d'encadrer la probabilité de réalisation de cet événement A .

Dans les paragraphes suivants, nous allons présenter les concepts de base de cette théorie tels que la distribution des possibilités, la mesure de possibilité, la mesure de nécessité, les opérateurs de fusion et les critères possibilistes de décision. Ces concepts seront à la base du système d'interprétation possibiliste que nous allons décrire dans les chapitres suivants de ce manuscrit.

2.3.4.1 Distribution de possibilités

Considérons un univers composé de N singletons $\Omega = \{x_1, x_2, \dots, x_N\}$ et supposons que l'on se trouve dans un contexte d'incertitude (c'est-à-dire qu'un seul singleton de Ω se produit mais nous ne le connaissons pas). La théorie des possibilités repose sur la notion de distribution de possibilités, notée $\pi(\cdot)$, attribuant à chaque singleton x_n de Ω une valeur dans l'intervalle $[0,1]$ qui encode notre état de connaissance ou croyance, sur la possibilité de la réalisation éventuelle de chaque singleton x_n de Ω . La valeur $\pi(x_n)$ encapsule nos connaissances liées à l'occurrence du singleton x_n . En d'autres termes, $\pi(x_n)$ représente dans quelle mesure il est possible que le singleton x_n soit l'unique singleton qui s'est produit.

Dans ce contexte, deux cas extrêmes des connaissances sont donnés :

- *Connaissance complète* : $\exists! x_n \in \Omega, \pi(x_n)=1$ et $\pi(x_m)=0, \forall x_m \in \Omega, x_m \neq x_n$;
- *Ignorance totale* : $\forall x_n \in \Omega, \pi(x_n)=1$ (tous les singletons sont considérés comme tout à fait possibles).

Dans le cas d'une connaissance complète, seul un singleton dispose de la valeur maximale de possibilité tandis que tous les autres singletons ont une possibilité d'occurrence nulle. En revanche, le cas de l'ignorance totale désigne la situation où nous ne disposons d'aucune connaissance permettant de favoriser (ou défavoriser) les différents singletons ; par conséquent, tous les singletons se voient attribuer le même niveau maximal de possibilité.

La valeur de possibilité maximale, $h(\pi) = \max_{x_n \in \Omega} (\pi(x_n))$, appelée la *hauteur* de la distribution de possibilités, est un indicateur qui mesure le *degré de consistance* (ou compatibilité) avec les connaissances disponibles (représentées par la distribution de possibilités π) concernant la réalisation des singletons. Cet indicateur détermine le (ou les) singleton(s) ayant le degré de compatibilité le plus élevé avec les connaissances disponibles.

La distribution $\pi(\cdot)$ est appelée distribution de possibilités *normale* (elle est dite *consistante* avec les connaissances disponibles) s'il existe au moins un singleton $x_n \in \Omega$ qui est totalement possible (i.e. $h(\pi)=1$). Dans le cas contraire (i.e. $h(\pi) < 1$), cette distribution est dite *non normalisée* ou *inconsistante*.

avec les connaissances disponibles et, dans ce cas, un nouvel indicateur $Inc(\pi) \in [0, 1]$ peut être utilisé pour caractériser ce degré d'inconsistance de la distribution de possibilités π :

$$Inc(\pi) = \max_{x_n \in \Omega} (\pi(x_n)) = 1 - h(\pi) \quad (2.12)$$

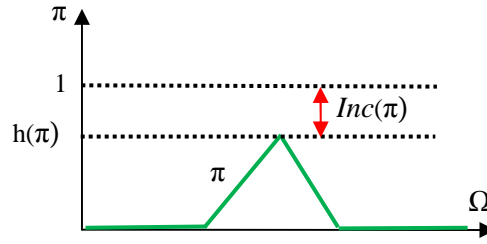


Figure 2.7 : Distribution de possibilités non-normalisée et degré d'inconsistance

2.3.4.2 Mesure de possibilité et mesure de nécessité

En se basant sur le concept de distribution de possibilités, deux mesures sont dérivées pour caractériser l'incertitude liée à l'occurrence de chaque sous-ensemble (aussi appelé un *événement*) $A \subseteq \Omega$: la *mesure de possibilité* Π et la *mesure de nécessité* N . Ces deux mesures sont définies par :

$$\begin{aligned} \Pi(A) &= \max_{x_n \in A} (\pi(x_n)) \\ N(A) &= 1 - \Pi(A^c) \end{aligned} \quad (2.13)$$

où A^c est l'événement complémentaire de A (i.e. $A \cup A^c = \Omega$ et $A \cap A^c = \emptyset$).

La mesure de possibilité $\Pi(A)$ estime le niveau de consistance (avec les connaissances disponibles représentées par la distribution de possibilités π) de la réalisation de l'événement A . Elle représente dans quelle mesure il est possible que l'unique singleton x_n de Ω qui s'est produit, soit dans le sous-ensemble A . Si $\Pi(A)=1$ (resp. $\Pi(A) = 0$), l'événement A est considéré comme un événement totalement possible (resp. impossible) à s'être produit (i.e. l'unique singleton x_n de Ω qui s'est produit est totalement possible qu'il soit dans A).

Notons que cette mesure de possibilité satisfait les propriétés suivantes :

- $\Pi(\emptyset)=0$ et $\Pi(\Omega)=1$;
- $\forall A, B \subseteq \Omega, \Pi(A \cup B) = \max(\Pi(A), \Pi(B))$;
- $\forall A, B \subseteq \Omega, \Pi(A \cap B) \leq \min(\Pi(A), \Pi(B))$.

En revanche, la mesure de nécessité $N(A)$ représente le niveau de certitude sur la réalisation de l'événement A , impliqué par les connaissances disponibles (représentées par la distribution de possibilités π). Elle représente dans quelle mesure on est certain que l'unique singleton x_n de Ω qui s'est produit, soit dans le sous-ensemble A . Si $N(A) = 1$ (resp. $N(A) = 0$), la réalisation de l'événement A est considérée comme totalement certaine (resp. incertaine) de s'être produite.

La mesure de nécessité satisfait les propriétés suivantes :

- $N(\emptyset)=0$ et $N(\Omega)=1$;
- $\forall A, B \subseteq \Omega, N(A \cup B) \geq \max(N(A), N(B))$
- $\forall A, B \subseteq \Omega, N(A \cap B) = \min(N(A), N(B))$

Ainsi, l'incertitude liée à la réalisation de chaque événement $A \subseteq \Omega$ est caractérisée par l'intervalle possibiliste $[N(A), \Pi(A)]$ qui permet de délimiter la probabilité de réalisation de A (i.e. $N(A) \leq \Pr(A) \leq \Pi(A)$) [78].

Notons aussi les propriétés suivantes concernant ces deux mesures qui permettent de distinguer clairement des mesures probabilistes :

- $\Pi(A) + \Pi(A^C) \geq 1$;
- $N(A) + N(A^C) \leq 1$;
- $\min\{N(A), N(A^C)\} = 0$;
- $\max\{\Pi(A), \Pi(A^C)\} = 1$;
- $\Pi(A) < 1 \Rightarrow N(A) = 0$;
- $N(A) > 0 \Rightarrow \Pi(A) = 1$

Ces propriétés sont essentielles pour comprendre le comportement décisionnel dans un environnement possibiliste (comme c'est le cas dans notre étude d'interprétation de scène).

2.3.5 Règles de décision possibiliste

Malgré le fait que la théorie des possibilités soit utilisée pour gérer l'incertitude, ce qui signifie qu'un contenu informationnel unique de Ω (qui est inconnu) se produise et que l'objectif ultime consiste à déterminer ce contenu, il est étonnant de constater que la question de la prise de décision est très peu abordée dans cette théorie. Dans ce paragraphe, nous allons étudier les règles existantes de prise de décision dans un cadre possibiliste.

Supposons que les connaissances disponibles sur la réalisation des éléments de l'univers Ω soient de nature ambiguë. Elle sont donc représentées par une distribution de possibilités associant à chaque singleton x_n une valeur de possibilité $\pi(x_n)$ représentant la possibilité pour que x_n soit l'unique singleton qui s'est produit.

La problématique de prise de décision possibiliste consiste à définir une règle de décision permettant d'identifier, à partir de la connaissance de la distribution $\pi(\cdot)$, le singleton x_n le plus vraisemblable. Etant donné que les deux mesures ensemblistes de possibilité $\Pi(A)$, et de nécessité $N(A)$, permettent de caractériser l'incertitude liée à la réalisation de chaque événement $A \subseteq \Omega$, une décision possibiliste peut être prise en se basant sur ces deux mesures. Une fois ces deux mesures estimées pour tous les événements de Ω , une règle de décision peut être déterminée pour connaître l'événement qui s'est produit c'est-à-dire l'événement qui contient l'unique singleton x_n qui s'est produit.

Pour le processus de prise de décision, nous nous limitons aux événements simples notés $A_n = \{x_n\}$, $n=1, 2, \dots, N$, qui ne comportent qu'un seul singleton x_n .

Deux familles de règles peuvent être distinguées : les règles basées sur l'utilisation directe de l'information encapsulée dans la distribution de possibilités, et les règles basées sur l'utilisation des mesures d'incertitude associées à cette distribution de possibilités.

2.3.5.1 Règle de décision basée sur la possibilité maximale

La règle de décision basée sur la possibilité maximale est sans doute la règle de décision la plus « intuitive » et la plus utilisée en classification possibiliste. Elle consiste à sélectionner l'événement A_{n_0} ayant la mesure de possibilité la plus élevée :

$$(R1) : \text{Décision } A_{n_0} \text{ ssi } \forall A_n \subseteq \Omega, n \neq n_0, \Pi(A_n) < \Pi(A_{n_0}) \quad (2.14)$$

Etant donné que la mesure de possibilité d'un événements $A_n = \{x_n\}$ se réduit à la valeur de possibilité du singleton (i.e. $\Pi(A_n) = \pi(x_n)$), cette règle consiste à sélectionner l'événement A_{n_0} qui contient le singleton ayant le degré de compatibilité le plus élevé avec les connaissances disponibles (i.e. $\pi(x_{n_0}) = \max_{x_n \in \Omega} (\pi(x_n))$).

2.3.5.2 Règle de décision basée sur la mesure de nécessité maximale

Cette règle de décision est réalisée en fonction de la mesure de nécessité maximale seule (i.e. en fonction de la borne inférieure de la probabilité d'occurrence). Elle consiste à sélectionner l'événement A_{n_0} ayant la mesure de nécessité la plus élevée.

$$(R2) : \text{Décision } A_{n_0} \text{ ssi } \forall A_n \subseteq \Omega, n \neq n_0, N(A_n) < N(A_{n_0}) \quad (2.15)$$

Dans [131], l'équivalence entre cette règle de décision avec la règle de décision basée sur la possibilité maximale a été donnée. Par conséquent, cette règle décisionnelle n'apporte aucun avantage supplémentaire par rapport à la règle de décision basée sur la possibilité maximale.

2.3.5.3 Règle de décision basée sur la maximisation de l'indice de confiance

D'autres règles de décision possibilistes sont développées sur la base des mesures d'incertitude. Parmi ces règles, on peut citer la règle basée sur l'indice de confiance proposé par S. Kikuchi *et al* [76].

En effet, l'indice de confiance est une mesure d'incertitude définie comme une combinaison des mesures de possibilité et de nécessité afin de quantifier l'incertitude liée à la réalisation de chaque événement $A \subseteq \Omega$:

$$\begin{aligned} Ind : 2^\Omega &\rightarrow [-1, 1] \\ A \rightarrow Ind(A) &= \Pi(A) + N(A) - 1, \quad \forall A \subseteq \Omega \end{aligned} \quad (2.16)$$

Cet indice présente les propriétés suivantes :

- $Ind(A) \in [-1, 1]$;
- $Ind(A) = -1$, si $N(A) = \Pi(A) = 0$ (ce cas désigne une totale invalidation de la réalisation de A);
et
- $Ind(A) = +1$, si $N(A) = \Pi(A) = 1$ (ce cas désigne la totale validation de la réalisation de A).

Cette règle de décision consiste à sélectionner l'événement A_0 ayant l'indice de confiance le plus élevé :

$$(R3) : \text{Décision } A_{n_0} \text{ ssi } \forall A_n \subseteq \Omega, n \neq n_0, Ind(A_n) < Ind(A_{n_0}) \quad (2.17)$$

En se limitant (pour le processus décisionnel) aux événements A_n , l'indice de confiance peut s'écrire de la manière suivante :

$$\begin{aligned} Ind(A_n) &= \Pi(A_n) + N(A_n) - 1 \\ &= \pi(x_n) + 1 - \Pi(A_n^c) - 1 \\ &= \pi(x_n) + 1 - \max_{k \neq n} \pi(x_k) - 1 \\ &= \pi(x_n) - \max_{k \neq n} \pi(x_k) \end{aligned}$$

$Ind(A_n)$ mesure donc l'écart entre la valeur de possibilité du singleton x_n et la valeur maximale de possibilité des autres singletons de Ω . Par conséquent, et ceci dans les deux cas de distributions normalisées ou non, seul l'élément ayant la valeur de possibilité maximale produira une valeur d'indice de confiance **strictement** positive, et tous les autres singletons produiront des valeurs négatives de cet indice (Figure 2.8).

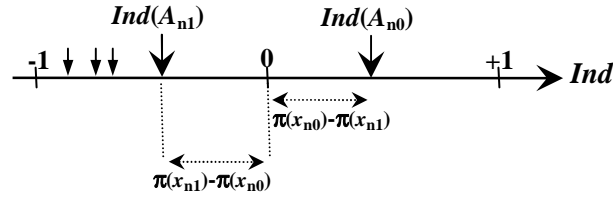


Figure 2.8 : Indices de confiance associés aux différentes décisions (A_{n0} : décision ayant la possibilité maximale, A_{n1} : décision ayant la deuxième plus grande possibilité)

Dans le cas où plusieurs éléments présentent la même valeur maximale de possibilité, cet indice prend la valeur zéro. On en déduit donc, que cette règle est équivalente à la règle de décision basée sur la possibilité maximale. Le véritable intérêt de cette règle de décision réside dans sa version la plus stricte qui n'accepte la prise de décision que si la valeur maximale de l'indice de confiance $Ind(A_{n0})$ dépasse un certain seuil S prédéfini (que l'on appelle *seuil possibiliste de confiance*) autorisant ainsi, l'option de rejet. Cette règle est donc formalisée par :

$$\begin{aligned}
 \text{(R3-rejet) : Décision } A_{n0} \text{ ssi, } \forall A_n \subseteq \Omega, n \neq n_0, Ind(A_{n0}) > Ind(A_n), \text{ et} \\
 Ind(A_{n0}) \geq S \quad \quad \quad (2.18) \\
 \text{Sinon} \quad \text{Décision} = \text{Rejet}
 \end{aligned}$$

Notons que le seuil possibiliste de confiance S , pour la règle (R3-Rejet), est basé sur l'écart entre les deux plus grandes valeurs de possibilités associées aux singletons de Ω , et qu'aucune décision n'est prise si la certitude liée à la réalisation de l'événement A_{n0} ne dépasse pas le seuil S .

2.4 Fusion d'information dans un cadre possibiliste

La fusion des informations est le processus qui consiste à combiner les informations issues de différentes sources de connaissances afin d'améliorer les performances des traitements et donc, l'interprétation de ces informations [71]. La fusion est particulièrement intéressante dans le cas d'informations imparfaites et hétérogènes afin d'obtenir des informations plus fiables qui permettent une meilleure prise de décision. L'exploitation conjointe de plusieurs sources de connaissances permet, grâce à la complémentarité et à la redondance des informations disponibles, de faire face à l'imperfection de ces sources.

Il est bien évident que le besoin de la fusion de différentes informations imparfaites est crucial dans le cadre des systèmes d'interprétation de scène. C'est pour cette raison que nous allons décrire brièvement ces opérateurs de fusion d'informations possibilistes.

La théorie des possibilités propose plusieurs familles d'opérateurs de combinaison d'informations issues de plusieurs sources possibilistes [72][11][75]. La capacité d'évaluer ces informations ou plutôt ces sources en termes de fiabilité, complémentarité, redondance et conflit permet de choisir l'opérateur de fusion le plus adapté aux spécificités des informations.

Soient π_1 et π_2 deux distributions de possibilités représentant des informations fournies respectivement par deux sources S_1 et S_2 . Nous allons présenter les principales familles d'opérateurs de fusion : les opérateurs conjonctifs, disjonctifs, et adaptatifs.

2.4.1 Opérateurs de fusion conjonctive

Les opérateurs conjonctifs (ou opérateurs d'intersection ensembliste T-norme) notés \wedge , sont utilisés pour la combinaison des sources fiables qui ne sont pas en contradiction totale (i.e. l'intersection des

supports des distributions de possibilités n'est pas l'ensemble vide). Ces sources sont dites « *concordantes* » et elles fournissent des informations qui sont en accord. La distribution de possibilités finale d'un opérateur conjonctif représente les informations fournies par les deux sources conjointement. Par conséquent, l'effet d'un opérateur conjonctif est de réduire l'incertitude globale en permettant de produire l'information la plus prometteuse, c'est-à-dire celle pour laquelle toutes les sources s'accordent à dire qu'elle devrait être la bonne valeur que l'on cherche. Différents opérateurs conjonctifs ont été proposés [73]:

- Le minimum (Zadeh T-norme) : $\pi_{\wedge}(x) = \pi_1(x) \wedge \pi_2(x) = \min(\pi_1(x), \pi_2(x))$;
- Le produit probabiliste : $\pi_{\wedge}(x) = \pi_1(x) \wedge \pi_2(x) = \pi_1(x) \cdot \pi_2(x)$;
- Le produit linéaire (Lukasiewicz T-norme) : $\pi_{\wedge}(x) = \pi_1(x) \wedge \pi_2(x) = \max(0, \pi_1(x) + \pi_2(x) - 1)$;
- etc.

La Figure 2.9 montre un exemple de l'utilisation de l'opérateur conjonctif « *min* » pour fusionner deux distributions de possibilités π_1 et π_2 .

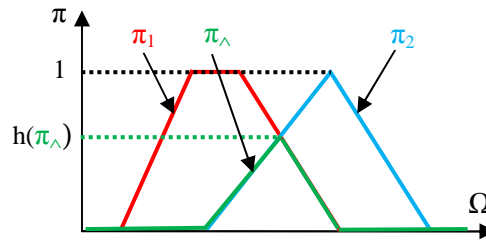


Figure 2.9 : Opérateur de fusion conjonctive : *min*

Notons que la hauteur de la distribution résultante $h(\pi_{\wedge}) = h(\pi_1, \pi_2) = \max(\min(\pi_1(x), \pi_2(x)))$, constitue un indicateur mesurant le degré de consistance (ou degré de cohérence) entre les deux distributions de possibilités. Dans ce contexte, deux cas extrêmes de consistance sont donnés :

- les deux sources sont dites en *total accord* : $h(\pi_{\wedge}) = 1$, c'est-à-dire $\exists x_n \in \Omega, \pi_1(x_n) = \pi_2(x_n) = 1$;
- les deux sources sont en *totale contradiction* : $h(\pi_{\wedge}) = 0$ (les deux distributions π_1 et π_2 n'ont aucune intersection informative).

En effet, plus les sources sont en conflit ou en contradiction partielle : $h(\pi_{\wedge}) < 1$, plus la fiabilité des sources est mise en cause. Pour faire face à ce type de situations, soit on élimine les sources non fiables, soit on prend en compte des informations supplémentaires [73].

2.4.2 Opérateurs de fusion disjonctive

Les opérateurs disjonctifs (ou opérateurs d'union ensembliste, T-conorme) notés \vee , sont utilisés pour la combinaison entre sources discordantes (i.e. en partielle ou totale contradiction). Ces opérateurs ont été conçus de telle sorte qu'ils gardent autant d'information que possible des deux distributions de possibilités parce que l'information sur l'unique singleton x que l'on cherche est certainement encodée par au moins l'une des sources (au moins l'une des sources disponibles est supposée être fiable) mais on ne sait pas laquelle. La distribution de possibilités finale d'un opérateur disjonctif représente toutes les informations fournies par les différentes sources. Cet opérateur augmente donc l'incertitude globale et fournit un résultat fiable mais moins précis que chacune des sources prise séparément. Ceci réduit sans doute, son utilité.

Différentes formes d'opérateurs disjonctifs ont été proposées [73]:

- Le maximum (Zadeh T-conorme) : $\pi_{\vee}(x) = \pi_1(x) \vee \pi_2(x) = \max(\pi_1(x), \pi_2(x))$;

- La somme probabiliste : $\pi_{\vee}(x) = \pi_1(x) \vee \pi_2(x) = \pi_1(x) + \pi_2(x) - \pi_1(x) \cdot \pi_2(x)$;
- Le produit linéaire (Lukasiewicz T-conorme) : $\pi_{\vee}(x) = \pi_1(x) \vee \pi_2(x) = \min(\pi_1(x) + \pi_2(x), 1)$;
- etc.

La Figure 2.10 montre un exemple de l'opérateur conjonctif « max » pour fusionner les deux distributions de possibilités π_1 et π_2 .

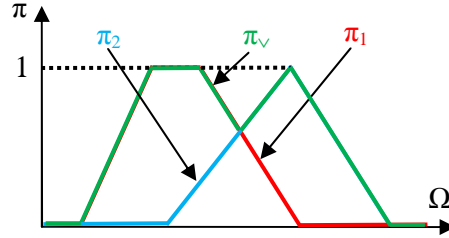


Figure 2.10 : Opérateur de fusion disjonctive : max

2.4.3 Opérateurs de fusion adaptative

Les opérateurs adaptatifs sont utilisés s'il n'y pas un consensus au niveau de la fiabilité des différentes sources (i.e. les sources d'information souffrent d'un conflit partiel). Dans ce cas, les opérateurs adaptatifs sont proposés afin préserver et renforcer l'information consensuelle, et de préserver, à moindre échelle, l'information conflictuelle entre les sources d'information. Un exemple d'opérateur adaptatif a été proposé par Dubois et Prade [73]. Cet opérateur permet de passer continûment du mode conjonctif au mode disjonctif dès que le conflit entre les deux sources augmente :

$$\pi_{DP}(x) = \max [\pi_{\wedge}(x)/h(\pi_1, \pi_2), \min(\pi_{\vee}(x), 1 - h(\pi_1, \pi_2))] \quad (2.19)$$

La Figure 2.11 montre l'opérateur adaptatif pour fusionner les deux distributions de possibilités π_1 et π_2 . En effet, cette opération est construite pour renforcer, le comportement du mode conjonctif (le pic du milieu) par rapport au comportement du mode disjonctif.

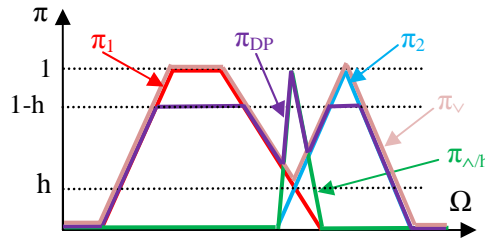


Figure 2.11 : Opérateur de fusion adaptative de Dubois et Prade

2.5 Théorie des possibilités en interprétation de scène

Dans les différents niveaux sémantiques d'un processus d'interprétation de scène (sous-pixelique, pixelique, région, objet ou scène), la présence d'informations imparfaites de nature possibiliste (i.e. incertaines et décrites avec des connaissances épistémiques) peut être facilement constatée. Ces connaissances épistémiques trouvent leurs origines soit dans la description réalisée par les experts de leurs connaissances par le biais de variables linguistiques, soit dans le fait d'utiliser un faible nombre d'échantillons ce qui rend non fiable la modélisation de ces échantillons (caractéristiques) par des distributions de probabilités. De plus, la coopération conjointe ou la fusion, est indispensable entre des mesures numériques (issues du système d'imagerie) et aussi certaines connaissances *a priori* sur la scène

exprimées par les experts de manière vague ou ambiguë. Cette coopération vise à identifier le contenu sémantique de la scène observée.

Nous constatons donc, que dans le cadre des systèmes d'interprétation de scène, la théorie des possibilités représente un cadre logique et légitime pour la représentation des connaissances et pour le traitement de ces connaissances. Cette théorie possède plusieurs avantages qui la rendent très utile dans un processus d'interprétation de scène :

- sa capacité de représentation des connaissances ambiguës ;
- les outils simples et pratiques proposés par cette théorie pour le traitement et la fusion des informations ambiguës et pour la quantification de l'incertitude liée à l'occurrence des événements de l'ensemble des contenus informationnels (en utilisant les mesures de possibilité et de nécessité) ;
- la simplicité et la flexibilité (en termes de calcul) pour le traitement et la fusion des informations issues de différentes sources de connaissances.

De ce fait, les chercheurs ont de plus en plus recours à l'utilisation de cette théorie pour la modélisation des différents niveaux sémantiques d'interprétation de scène [3][7][29][37][111]. Néanmoins, l'analyse effectuée dans le premier chapitre (concernant les approches possibilistes) montre certains inconvénients liés à l'utilisation de la théorie des possibilités pour l'interprétation de scène.

En effet, l'estimation des distributions de possibilités caractérisant les différentes classes dans une scène constitue l'un des inconvénients majeurs rencontrés. Généralement, cette estimation est basée sur la description par l'expert des différentes classes, donnée sous la forme de termes linguistiques sur la base des formes de distributions de possibilités prédéfinies (pour représenter cette description). Ceci nécessite l'estimation des paramètres de ces formes tout en offrant un cadre limité de formes de distributions de possibilités. Dans un contexte où les classes sont très difficiles à discriminer (images très bruitées, fortement texturées, classes multimodales,...), cette description devient très difficile, voire peu fiable car nous ne disposons que de zones « sur-représentatives » des classes thématiques présentes dans la scène (du point de vue de l'expert). L'estimation des distributions de possibilités caractérisant les différentes classes à partir de ces zones constitue donc un problème récurrent dans l'utilisation de la théorie des possibilités.

Un autre inconvénient des approches possibilistes proposées jusqu'à présent pour l'interprétation de scène, réside dans le fait que ces approches accordent peu d'importance au contexte spatial et à son rôle pour réduire l'ambiguïté liée à la caractérisation des classes thématiques. Rappelons à ce stade, que le contexte spatial a été exploité dans l'approche proposée dans [111] afin de réduire l'incertitude liée à la classification des pixels. Dans le travail que nous proposons, plusieurs contributions sont proposées afin d'apporter certaines réponses aux lacunes et questions encore ouvertes que nous avons soulevées, dans un contexte d'application de la théorie des possibilités en interprétation de scène.

2.6 Conclusion

Dans ce chapitre nous avons présenté la caractérisation de l'information et les différentes formes d'imperfections pouvant l'entacher tant au niveau intrinsèque qu'au niveau contextuel. La représentation de ce type d'information via un modèle mathématique doit simplifier la fusion d'informations en vue d'améliorer l'interprétation d'une scène observée. Les approches essentielles pour la représentation des informations imparfaites ont été présentées brièvement : la théorie des probabilités, la théorie des croyances, la théorie des ensembles flous et la théorie des possibilités.

Nous avons apporté une attention particulière à la théorie des possibilités qui représente un cadre prometteur pour la représentation des informations entachées d'incertitude et décrites par des connaissances ambiguës (i.e. les informations possibilistes). La notion de distribution de possibilités constitue l'outil essentiel de la théorie des possibilités pour représenter les connaissances disponibles. Deux mesures ensemblistes dérivées de cette distribution ont été présentées : la mesure de possibilité (resp. de nécessité) qui indique la compatibilité (resp. la certitude) d'occurrence de chaque sous-ensemble (ou événement) A du contenu informationnel Ω . La combinaison de ces deux mesures définit un intervalle possibiliste permettant d'avoir une quantification de l'incertitude liée à l'occurrence de chaque événement A . Nous avons ensuite présenté les différents opérateurs de fusion d'informations possibilistes. A la fin de ce chapitre, nous avons présenté l'intérêt et l'adéquation de l'utilisation de la théorie des possibilités dans un processus d'interprétation de scène tout en mettant en avant les principaux avantages de cette théorie.

Cependant, il est indispensable d'aborder la question de l'estimation des distributions de possibilités à partir de zones limitées d'échantillons et ceci dans un contexte où les classes thématiques contenues dans la scène en cours d'analyse sont difficiles à discriminer.

Chapitre 3

ESTIMATION DES DISTRIBUTIONS DE POSSIBILITES DANS LES SYSTEMES DE CLASSIFICATION PIXELIQUE

3.1 Introduction

L'objectif de notre travail de recherche consiste à proposer une approche d'analyse hiérarchique/sémantique de scène basée sur l'application de la théorie des possibilités. Cette analyse considère les différents niveaux sémantiques d'une scène (i.e. les niveaux : sous-pixelique, pixelique, région, objet et scène) et propose des solutions à certaines problématiques associées à ces niveaux (démixage, segmentation, classification, etc.) selon différentes démarches ascendante ou descendante, de propagation des connaissances et des résultats issus de cette analyse dans les différents niveaux sémantiques.

La théorie de possibilités introduite par Zadeh en 1978 et brièvement présentée dans le chapitre précédent, est considérée à la base de notre travail pour différentes raisons. En effet, cette théorie de par sa flexibilité, sa capacité à réaliser la représentation des connaissances imparfaites et les outils de raisonnement qu'elle propose, permet de faire face à la situation d'incertitude en prise de décision où les connaissances disponibles sont de nature ambiguë [52]. Comme nous l'avons déjà précisé, ce type de situations est très fréquemment rencontré lorsque les connaissances d'experts sont exprimées en langage naturel via des variables linguistiques définissant des ensembles flous. Dans une telle situation, l'application de l'approche bayésienne (i.e. l'approche optimale de traitement de l'incertitude) se trouve démunie de justification physique (à cause des connaissances ambiguës et non probabilistes, disponibles). Il en est de même lorsque les données statistiques disponibles ne sont pas représentatives de la situation d'incertitude. C'est précisément dans ce contexte d'application que la théorie des possibilités a été proposée par L. Zadeh, D. Dubois, H. Prade et bien d'autres chercheurs.

Comme nous l'avons décrit dans le chapitre précédent, l'application de cette théorie consiste à encoder les connaissances ambiguës sous la forme d'une distribution de possibilités. Sur la base de cette distribution, un ensemble d'outils de raisonnement est alors disponible afin de fusionner les différentes connaissances ambiguës et former deux mesures ensemblistes (de possibilité et de nécessité), sur lesquelles des outils d'aide à la décision peuvent opérer.

Nous remarquons la dimension primordiale de la notion de distribution de possibilités. Néanmoins, l'analyse de la littérature liée à cette thématique montre une certaine faiblesse pour la détermination des distributions de possibilités face à l'ensemble impressionnant de travaux consacrés à l'estimation des densités de probabilités dans le contexte du raisonnement bayésien classiquement utilisé pour le traitement de l'incertitude.

Il est important de noter que l'estimation des distributions de possibilités souffre de plusieurs difficultés. Une première raison à ces difficultés réside en l'absence de règles permettant d'évaluer la qualité ou l'exactitude d'une telle distribution estimée à l'aide d'une méthode particulière [77]. Un autre problème qui rend l'estimation des distributions de possibilités non-triviale, est l'absence de consensus sur une signification (ou interprétation) unique « du sens physique » de ces distributions de possibilités. Par exemple, que signifie que John appartient à la classe des hommes « *grands* » avec la valeur d'appartenance 0.7 ? Dubois et Prade [78] ont donné trois interprétations possibles à cette valeur d'appartenance : (a) 70% de la population a déclaré que John est grand (*vision de vraisemblance*), (b) 70% de la population décrit « *grand* » comme un intervalle contenant la hauteur de John (*vision d'ensemble aléatoire*), et (c) la hauteur de John est à une distance normalisée égale à 0.3 à partir du prototype idéal de l'homme "grand" (*vision de compatibilité*). Par conséquent, plusieurs méthodes d'estimation des distributions de possibilités ont été proposées dans la littérature et ceci en fonction de l'interprétation de ces distributions.

Dans ce chapitre, nous nous focalisons sur cette problématique d'estimation de distributions de possibilités à partir d'un ensemble réduit de connaissances fournies par un expert. Plus précisément, et comme nous considérons le cadre d'analyse de scène à partir de sources d'acquisition d'informations du type image(s), cet ensemble de connaissances est soit formulé par « une forme approximative » de distributions de possibilités, ou bien par la délimitation de zones de faibles tailles caractérisant les classes thématiques présentes dans la scène analysée.

Pour notre contexte d'étude, nous supposons que nous disposons d'une ou de plusieurs images d'une scène comportant M classes thématiques C_m , $m = 1, 2, \dots, M$. Dans chaque image de la scène, le contenu informationnel de chaque classe est caractérisé par une ou plusieurs propriétés décrivant soit l'information brute en sortie du capteur (e.g. intensité lumineuse des pixels), soit une primitive de plus haut niveau (liée à la texture, le gradient de l'intensité, le contexte spatial, etc.). De plus, nous supposons que les connaissances *a priori* concernant chacune de ces propriétés de classes sont explicitées par un expert via des variables linguistiques (par exemple, les pixels sont *brillants* en vue de caractériser la zone de tumeur dans une image médicale) ou via la délimitation sur l'image de la scène étudiée, de zones comportant des pixels vérifiant les propriétés caractérisant la classe concernée. Il s'agit d'un contexte d'incertitude où les connaissances disponibles sont affectées par l'ambiguïté.

C'est dans ce contexte d'information imparfaite que nous proposons d'exploiter le contexte spatial afin de « raffiner » l'estimation des distributions de possibilités permettant de caractériser les différentes classes thématiques présentes dans une scène analysée.

Dans la suite de ce chapitre, nous nous concentrerons, dans un premier temps, sur les méthodes d'estimation des distributions de possibilités dans le domaine de l'interprétation de scène et plus spécialement, sur les méthodes proposées dans la littérature, dans le contexte de l'imagerie.

Ensuite, nous proposerons une méthode itérative de raffinement de l'estimation des distributions de possibilités. Sur la base d'une estimation initiale des distributions de possibilités (liée aux « faibles » connaissances exprimées par un expert), l'approche proposée consiste à exploiter le contexte spatial d'une façon itérative afin de « renforcer » les connaissances disponibles et de réaliser ainsi, un raffinement de l'estimation des distributions de possibilités. En d'autres termes, la méthode proposée ne constitue pas une nouvelle approche d'estimation des distributions de possibilités. En revanche, cette méthode doit être positionnée sur le plan d'une démarche méthodologique basée sur un raisonnement possibiliste spatial et permettant d'enrichir les connaissances exprimées par l'expert. Une étude expérimentale de la méthode proposée est ensuite menée en l'appliquant sur des images de synthèse ainsi que sur des images réelles afin d'en réaliser des évaluations quantitatives et qualitatives.

3.2 Etat de l'art des méthodes d'estimation des distributions de possibilité

De nombreuses méthodes d'estimation de distributions de possibilités ont été proposées dans la littérature, en particulier dans le domaine de la reconnaissance de formes [77]. On peut regrouper ces méthodes en trois catégories.

La première catégorie regroupe les méthodes d'estimation guidées par l'expert (i.e. les formes des distributions de possibilités sont sélectionnées par l'expert [80] [81] [82] [83]).

La deuxième catégorie est celle des méthodes basées sur des approches statistiques qui consistent à réaliser dans un premier temps, une estimation des densités de probabilités, puis ensuite, à transformer ces densités de probabilités en des distributions de possibilités [90] [91] [92] [93].

La troisième catégorie est celle des méthodes basées sur l'apprentissage à partir de données statistiques extraites de l'image [95] [97] [121] [100]. Dans ce qui suit, on propose de décrire brièvement, ces trois catégories tout en mettant l'accent sur la problématique de caractérisation des classes thématiques, prédéfinies dans la scène, par le biais des distributions de possibilités estimées à partir d'un ensemble limité d'échantillons (i.e. connaissances fournies par un expert).

3.2.1 Méthodes d'estimation guidées par l'expert

En se basant sur les connaissances exprimées par l'expert, ces méthodes d'estimation consistent à attribuer aux différents singletons (pixels, région, ou objets), des valeurs de possibilités partielles représentant la véracité ou l'adéquation de ces singletons avec les différentes propriétés caractérisant les différentes classes. Les connaissances *a priori* (exprimées par l'expert) sont liées aux différentes classes et sont souvent représentées par le biais de fonctions d'appartenance à des ensembles flous ayant des formes standards de sorte que l'expert n'a qu'à sélectionner la forme qui lui « semble » la plus adéquate.

Il est bien évident que l'efficacité de telles méthodes dépend du niveau des connaissances de l'expert et de sa capacité à représenter les propriétés caractérisant les classes thématiques dans la scène analysée d'une manière permettant la mesure de la force de liaison entre les singletons (pixels, les régions ou les objets) avec les propriétés considérées.

Ces méthodes ont été utilisées avec succès dans des applications de vision par ordinateur pour décrire certaines relations spatiales comme « au-dessus » et « à la gauche de » [79][80].

Cette catégorie des méthodes d'estimation guidées par l'expert peut elle-même être divisée en deux sous-catégories : les *méthodes heuristiques* où les distributions de possibilités sont entièrement déterminées par l'expert et les *méthodes hybrides* où les distributions de possibilités sont partiellement

déterminées par l'expert (i.e. l'expert sélectionne la forme standard des distributions de possibilités, et l'utilisation d'approches d'optimisation permet de déterminer les paramètres associés à ces formes standards).

3.2.1.1 Méthodes heuristiques

Les méthodes heuristiques d'estimation dépendent complètement de l'expert tant au niveau du choix de la forme standard de la distribution de possibilités (dupliquant la fonction d'appartenance à un ensemble flou défini sur l'espace d'une primitive donnée), qu'au niveau de l'estimation des paramètres associés à cette forme [81] [82] [83]. En se basant sur ses connaissances, l'expert réalise cette tâche d'une façon manuelle, empirique et selon des critères totalement subjectifs qu'il impose aux degrés de possibilités pour qu'ils correspondent au mieux aux propriétés décrivant les classes à caractériser. Cette subjectivité rend ce type de méthodes d'estimation moins robuste surtout dans les applications où les paramètres doivent être ajustés jusqu'à ce que la performance finale soit acceptable [77].

3.2.1.2 Méthodes hybrides

Dans la plupart des applications, il est très difficile pour l'expert, d'estimer la distribution de possibilités en se basant sur ses seules connaissances.

Il doit donc recourir à des méthodes automatiques qui, en intégrant certaines informations statistiques résumant l'opinion d'un expert, peuvent estimer les paramètres associés aux distributions de possibilités.

Si nous limitons la forme sélectionnée à un choix spécifique (par exemple, triangulaire, trapézoïdal, etc.) alors le problème d'estimation se réduit à un problème d'optimisation d'un nombre limité de paramètres.

A titre d'exemple, et à partir de connaissances *a priori* (i.e. zones d'apprentissage), les différentes classes présentes dans une image peuvent être caractérisées. En effet, pour chaque classe C_m contenue dans l'image, une zone d'apprentissage Z_m peut être définie par l'expert. Cette zone correspond à une petite région de l'image que l'expert juge représentative d'une classe. L'ensemble des pixels situés dans cette zone d'apprentissage vérifient les propriétés caractérisant la classe concernée. Cet ensemble de pixels permet d'extraire certaines informations statistiques telles que la moyenne et l'écart type [104]. Etant données ces informations statistiques, la distribution des possibilités π_{C_m} caractérisant la classe C_m peut être construite. Par exemple, dans le cas de l'utilisation d'un modèle triangulaire opérant sur l'intensité des niveaux de gris (Figure 3.1), le point du sommet de ce modèle est celui correspondant à la moyenne « m_m » des niveaux de gris de la zone d'apprentissage Z_m . Le support de ce modèle « Δ_m » est considéré comme fonction de l'écart type σ_m des niveaux de gris. Dans [104], $\Delta_m = 3\sigma_m$ car cette forme représente assez fidèlement une distribution normale.

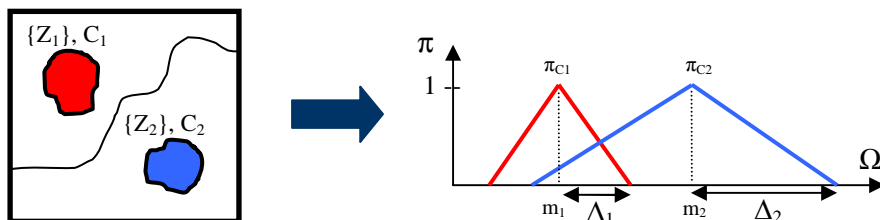


Figure 3.1 : Méthode à base d'estimation statistique des paramètres

Néanmoins, le processus d'optimisation peut être appliqué sur des mesures indicatives, calculées à partir de primitives et considérées comme des fonctions objectives à optimiser. Le calcul de ces mesures (telle que l'entropie floue ou l'indice flou [84] [85]) permet d'évaluer l'information ambiguë contenue

dans une classe. Ceci permet donc d'accomplir cette tâche d'estimation d'une manière plus efficace et surtout plus objective par rapport aux méthodes heuristiques et en termes d'optimisation des paramètres des distributions de possibilités. Dans ce sens, plusieurs algorithmes peuvent être employés pour l'optimisation des paramètres. On peut citer, à titre indicatif et non exclusif : le recuit simulé [86], les algorithmes génétiques [87] et l'algorithme d'optimisation par essais particuliers [88]. L'optimisation des paramètres des distributions de possibilités s'effectue, pour chacun de ces algorithmes, d'une manière stochastique [89]. Ceci explique le fait que les résultats obtenus par ces algorithmes peuvent différer d'une réalisation à une autre.

D'une manière générale, la caractérisation par des termes linguistiques des classes présentes dans une scène, devient très difficile dans un contexte de confusion interclasses où ces classes sont difficiles à discriminer (faible contraste, contexte très bruité, chevauchement spectral et spatial) ou tout "simplement" multimodales. Ainsi, dans ce contexte, le choix de la forme des distributions de possibilités constitue une difficulté majeure rencontrée lors de l'utilisation de telles méthodes d'estimation guidées par l'expert. Par conséquent, les méthodes basées sur apprentissage à partir de des données statistiques extraites de l'image, ou bien les méthodes basées sur la transformation d'une distribution de probabilités vers une distribution de possibilités peuvent être utilisées. Nous allons donc détailler ces méthodes.

3.2.2 Méthodes de transformation d'une distribution de probabilités en une distribution de possibilités

Une mesure de possibilité peut coder une famille de mesures de probabilité [91]. Ce fait est la base de l'utilisation de la transformation de l'information statistique/probabiliste en une information possibiliste qui permet d'avoir, pour chaque valeur de probabilité (ou valeur de fréquence d'occurrence), un intervalle possibiliste de confiance substituant cette valeur (ou un intervalle de compatibilité). Ainsi, une telle transformation permet de passer de l'espace probabiliste vers l'espace possibiliste tout en quantifiant l'incertitude liée à l'occurrence des singletons de Ω . En effet, dans plusieurs situations, les données statistiques disponibles souffrent d'imprécision. C'est le cas, par exemple, lorsque la taille de l'ensemble des données disponibles est si petite qu'elle ne peut pas constituer une représentation valide de l'information probabiliste. Une représentation possibiliste de l'information permet d'utiliser les mécanismes de raisonnement possibiliste afin d'extraire des nouvelles connaissances à partir de données disponibles. Une autre raison de l'existence d'une telle transformation est que la prise de décision dans un système de traitement d'information est généralement basée sur la fusion de plusieurs types d'informations qui sont affectées de différentes formes d'imperfection : information incertaine, information possibiliste, information ambiguë, etc. Afin de traiter ces différents types d'informations dans un cadre commun, une transformation d'un type de représentation de connaissances vers un autre devient incontournable [74].

Les méthodes de transformation d'une distribution de probabilités en une distribution de possibilités sont basées sur l'utilisation de données statistiques et sont réalisées en deux étapes consécutives :

En utilisant les données statistiques extraites d'une image ou d'une base d'apprentissage, une fonction d'incertitude (décrivant l'incertitude inhérente aux données statistiques) est d'abord estimée (par exemple, une densité de probabilité, une fonction de croyance, etc.). Cette fonction d'incertitude est ensuite transformée en une distribution de possibilités.

Dans de nombreuses applications, les histogrammes normalisés ont été traditionnellement considérés comme des estimations de densités de probabilités. Si on dispose d'un grand nombre d'échantillons appartenant à une région d'intérêt dans une image, l'histogramme normalisé de ces échantillons peut être considéré comme une approximation d'une densité de probabilités caractérisant la classe/région

particulière. A ce stade-là, les méthodes de transformation d'une distribution de probabilités en une distribution de possibilités peuvent être utilisées pour générer des distributions de possibilités à partir des histogrammes.

Établir une relation entre la représentation d'informations probabiliste et la représentation d'informations floue (plus largement possibiliste) a bénéficié d'une attention particulière depuis les propos de Zadeh à cet égard [52]. Plus récemment, ce problème a été étudié par plusieurs auteurs (par exemple, Dubois *et al.* [90] [91], Klir [92], Yamada [93], etc.) en se basant sur le principe de cohérence probabilités-possibilités proposé par Zadeh [52]. Dans la suite, nous allons présenter quelques principes reliant les distributions de probabilités à celles des possibilités, tout en détaillant la transformation probabilités-possibilités proposée par Dubois-Prade qui est adoptée dans notre travail de recherche à cause des bonnes performances de cette méthode de transformation (dans le domaine de la reconnaissance de formes [110]), ainsi que la possibilité de l'appliquer dans les deux directions de transformation.

3.2.2.1 Principe de cohérence probabilités-possibilités de Zadeh

Zadeh a défini le principe de cohérence probabilités-possibilités comme suit :

« Un fort degré de possibilité n'implique pas un degré élevé de probabilité. Un faible degré de probabilité n'implique pas non plus, un faible degré de possibilité. Toutefois, si un événement est impossible, il est forcément improbable » [52].

L. Zadeh a souligné que la cohérence probabilités-possibilités est une formalisation approximative de l'observation heuristique qui consiste à dire « qu'une diminution de la possibilité d'un événement tend à diminuer sa probabilité, mais pas vice-versa » [52].

Soit une variable aléatoire x prenant ses valeurs dans l'univers de discours $\Omega = \{x_1, x_2, \dots, x_N\}$ et considérons deux représentations de cette variable :

- Par une distribution de probabilités $\Pr(x_n)$, $n=1, 2, \dots, N$ et
- Par une distribution de possibilités $\pi(x_n)$.

Zadeh a défini le degré de cohérence C_z entre ces deux distributions comme:

$$C_z = \sum_{n=1}^N \Pr(x_n) \cdot \pi(x_n) \quad (3.1)$$

3.2.2.2 Transformation de Dubois et Prade

Dubois *et al.* [90] ont suggéré que toute transformation d'une distribution de probabilités \Pr en une distribution de possibilités π , devrait être guidée par les deux principes suivants :

- *Le principe de cohérence probabilités-possibilités de Zadeh.* Dubois et Prade [78] ont formulé ce principe en indiquant que la mesure de possibilité d'un événement $A \subseteq \Omega$ doit être supérieure ou égale à sa mesure de probabilité, qui doit être elle-même supérieure ou égale à sa mesure de nécessité :

$$N(A) \leq \Pr(A) \leq \Pi(A), \forall A \subseteq \Omega \quad (3.2)$$

$$\text{où } \Pi(A) = \max_{x_n \in A} \pi(x_n), \Pr(A) = \sum_{x_n \in A} \Pr(x_n) \text{ et } N(A) = 1 - \Pi(A^C) = \min_{x_n \notin A} \{1 - \pi(x_n)\}.$$

- *Le principe de préservation de préférence.* Ce principe fait en sorte que toute transformation d'une distribution de probabilités \Pr en une distribution de possibilités π doit satisfaire la relation suivante [91]:

$$\Pr(A) < \Pr(B) \Leftrightarrow \Pi(A) < \Pi(B), \quad \forall A, B \subseteq \Omega \quad (3.3)$$

En vérifiant ces deux principes, une transformation d'une distribution de probabilités \Pr en une distribution de possibilités π a été suggérée par Dubois et Prade [90]. Cette transformation est définie par :

$$\pi(x_n) = \sum_{j=1}^N \min[\Pr(x_j), \Pr(x_n)] \quad (3.4)$$

Cette transformation définie par l'Equation (3.4) est appelée *symétrique* car elle est applicable dans les deux directions de transformation $\Pr \rightarrow \pi$ et $\pi \rightarrow \Pr$. En effet, la transformation « inverse », donnée en Equation (3.5) et appliquée à la distribution de possibilités obtenue, permet de retrouver la distribution de probabilités initiale :

$$\Pr(x_n) = \sum_{j=1}^N \frac{\pi(x_j) - \pi(x_{j+1})}{j} \quad (3.5)$$

où $\pi(x_1) = 1 \geq \pi(x_2) \geq \dots \geq \pi(x_N)$ et $\pi(x_{N+1}) = 0$.

3.2.3 Méthodes à base d'apprentissage.

A la différence des méthodes présentées précédemment, les méthodes à base d'apprentissage adoptent une approche d'estimation directe des distributions de possibilités à partir de données statistiques. En analyse de scène, ces données statistiques peuvent être formées par des zones, délimitées par un expert sur l'image de la scène étudiée, comportant des pixels vérifiant les propriétés caractérisant chacune des classes prédéfinies.

Parmi les méthodes à base d'apprentissage, nous citons l'algorithme Fuzzy C-Means (FCM) et la méthode basée sur l'algorithme des k plus proches voisins. Ces deux méthodes sont largement utilisées pour l'estimation des distributions de possibilités à partir de données statistiques.

3.2.3.1 L'algorithme Fuzzy C-Means (FCM)

L'algorithme Fuzzy C-Means (FCM) [94] est l'une des techniques les plus utilisées pour la segmentation et la classification des images [95] [97]. Cette technique permet de générer une partition floue plutôt qu'une partition classique d'un ensemble fini de données. A l'issue de cette partition, des ensembles flous caractérisant les classes prédéfinies dans l'image, sont formés. Ainsi, chaque échantillon (i.e. pixel) de l'image peut appartenir à plusieurs classes, avec différentes valeurs d'appartenance à ces classes.

Considérons $S = \{P_1, P_2, \dots, P_N\}$ un ensemble fini formé de N pixels où chaque pixel a un niveau de gris $x_i \in \Omega$, et supposons M , le nombre de classes thématiques. L'algorithme FCM permet de déterminer un ensemble des centres de gravité des classes $V = \{v_1, v_2, \dots, v_M\}$ et une matrice $U = [\mu_{ij}] \in [0,1]$, $i = 1, \dots, N$, $j = 1, \dots, M$ avec $\sum_{j=1}^M \mu_{ij} = 1$, représentant la partition floue générée par cet algorithme. Les valeurs d'appartenance μ_{ij} sont estimées par l'algorithme FCM en minimisant d'une manière itérative, la fonction objective suivante :

$$\mu_{ik} = \frac{(\|v_k - x_i\|)^{2/(m-1)}}{\sum_{j=1}^M (\|v_j - x_i\|)^{2/(m-1)}} \quad (3.6)$$

où μ_{ik} représente le degré d'appartenance du pixel P (dont le niveau de gris est x_i) à une classe k et $\|v_j - x_i\|$ est la distance euclidienne entre le centre de la classe C_j et le niveau de gris du pixel P. Le paramètre $m \in [1, \infty[$ est un paramètre qui permet de fixer le niveau du « flou » de la partition obtenue. Dans le cas où $m=1$, les valeurs d'appartenance μ_{ij} convergent vers 0 ou 1 et donc chaque pixel est attribué à une classe donnée unique.

En raison de l'utilisation de la contrainte probabiliste représentée par $\sum_{j=1}^M \mu_{ij} = 1$, les degrés d'appartenance générées par l'algorithme FCM ne représentent pas la notion intuitive de degrés d'appartenance ou compatibilité, mais plutôt de degrés de partage ou de vraisemblance [96]. En d'autres termes, la génération du degré d'appartenance du pixel P dépend de tous les autres pixels dans l'image. Ceci ne permet pas à l'algorithme FCM de traiter les « outliers » et le rend très sensible au bruit [77]. Par ailleurs, l'algorithme FCM dépend de la notion de prototype (i.e. le centre de gravité de classe v) pour caractériser la forme des classes [101]. Ceci impose une forme des classes qui prend souvent une forme similaire à celle d'une gaussienne.

3.2.3.2 Méthode basée sur l'algorithme des k plus proches voisins

Etant donné un ensemble fini S formé de N pixels $S = \{P_1, P_2, \dots, P_N\}$ où chaque pixel a un niveau de gris $x_i \in \Omega$, l'algorithme des k plus proches voisins k-NN [98], est une méthode non paramétrique qui consiste à retenir la classe la plus représentée parmi les k pixels les plus proches du pixel de l'ensemble S , et ceci en utilisant la règle du vote majoritaire. La simplicité de mise en œuvre et les bons résultats obtenus ont rendu cette technique très populaire.

Keller *et al.* [99] ont introduit la notion d'ensemble flou à cet algorithme, aboutissant à l'algorithme Fuzzy k-NN qui permet d'attribuer à un pixel P d'une image I , des valeurs d'appartenance aux différentes classes prédéfinies dans la scène étudiée, en utilisant l'équation suivante :

$$\mu_i = \frac{\sum_{j=1}^k \mu_{ij} \cdot (1/\|x - x_j\|)^{2/(m-1)}}{\sum_{j=1}^k (1/\|x - x_j\|)^{2/(m-1)}} \quad (3.7)$$

où μ_{ij} représente le degré d'appartenance du pixel P_j de l'ensemble S à la classe C_i et $\|x - x_j\|$ est la distance euclidienne appliquée sur les intensités des pixels P et P_j .

Les méthodes floues basées sur l'algorithme k-NN n'imposent aucune allure à la fonction d'appartenance et elles ne dépendent que de la distance entre les échantillons de la classe considérée [99][100][101][102]. Cependant, et d'une façon semblable à l'algorithme FCM, l'algorithme k-NN souffre de l'utilisation de la contrainte probabiliste représentée par $\sum_{i=1}^M \mu_{ij} = 1$ [103]. En outre, cet algorithme ne génère pas de fonctions d'appartenance lisses dans les zones de chevauchement entre les classes [77].

Un inconvénient majeur des méthodes à base d'apprentissage réside dans le fait que la caractérisation des classes est réalisée à partir d'ensembles limités d'échantillons (données statistiques extraites à partir des zones représentatives, délimitées par un expert sur l'image). En effet, la génération des connaissances à partir de ces données limitées devient très difficile dans un contexte de confusion interclasses. Dans ce contexte, l'expert aura tendance à sélectionner les zones les plus représentatives et néglige les autres

zones qu'il considère moins représentatives, mais pourtant souvent majoritaires dans la scène ! De plus, la sélection des zones d'apprentissage par l'expert, reste empreinte de subjectivité, et donc peut différer d'un expert à l'autre. Par conséquent, la caractérisation des classes via des distributions de possibilités estimées juste à partir des zones d'apprentissage sélectionnées par l'expert devient une tâche très difficile.

Dans la suite de ce chapitre, nous proposons de présenter une approche permettant le raffinement des distributions de possibilités estimées à partir des zones d'apprentissage et ceci en vue de « mieux » caractériser les classes présentes dans la scène surtout dans un contexte où ces classes sont difficiles à discriminer.

3.3 Approche proposée pour le raffinement des distributions de possibilités

Dans les systèmes d'interprétation de scène, la caractérisation exprimée par l'expert des différentes classes, joue un rôle essentiel pour l'identification du contenu sémantique de cette scène. Cette caractérisation est réalisée, soit par le biais de termes linguistiques caractérisant les propriétés réalisées par les pixels appartenant à ces classes, soit par la délimitation manuelle (par l'expert) des zones comportant des pixels considérés comme caractéristiques des classes présentes dans la scène.

Notons que le choix des zones représentatives (vue de l'expert) des différentes classes constitue le moyen le plus simple (souvent adopté par l'expert) par rapport à la caractérisation par des termes linguistiques surtout dans un contexte où les classes sont difficiles à discriminer (confusion interclasses).

Il est bien évident que l'expert se focalise sur la délimitation des zones « les plus caractéristiques » des classes thématiques et évite ainsi les zones pouvant apporter une confusion pour la discrimination des classes. Ce qui signifie que l'expert a tendance à sélectionner les zones qu'il juge comme les plus pures et les moins bruitées des classes thématiques. Par conséquent, et d'un point de vue de l'ingénierie, la question de la représentativité de ces zones est fortement soulevée lorsque ces zones seront considérées pour l'apprentissage des connaissances de l'expert. Cette question de représentativité des zones caractéristiques reste toujours ouverte et aucune réponse générique n'a été apportée dans la littérature à cette question clef.

Dans ce paragraphe, nous proposons une approche de raffinement des distributions de possibilités estimées à partir de zones d'apprentissages délimitées par un expert et dont la représentativité des classes thématiques n'est pas forcément optimale. Cette approche exploite les connaissances issues de plusieurs sources de connaissances (description d'expert et contexte spatiale des pixels) et qui sont représentées dans un espace possibiliste.

En utilisant les mécanismes de raisonnement possibiliste spatial, l'approche proposée intègre de manière itérative, l'information issue du contexte spatial des pixels afin d'enrichir la base d'apprentissage initiale (raffinement) des différentes classes. En effet, l'exploitation du contexte spatial des pixels permet l'extraction de nouveaux pixels ayant une possibilité d'appartenance à une classe thématique donnée avec un fort degré de certitude, et ceci selon la règle de sélection (décision) adoptée.

Les pixels sélectionnés à chaque itération (appelés *germes possibilistes*) sont ajoutés à la base d'apprentissage des différentes classes. Notons que la base d'apprentissage initiale est formée par les zones d'apprentissage délimitées par l'expert.

Soit $\psi = \{C_1, C_2, \dots, C_M\}$ un ensemble exhaustif et exclusif de M classes thématiques contenues dans la scène analysée et soit $B_0 = \{Z_{0,1}, Z_{0,2}, \dots, Z_{0,M}\}$ un ensemble comportant des zones d'apprentissage

initiales pour les différentes classes thématiques prédéfinies, zones délimitées par l'expert dans une image I donnée de la scène analysée.

L'approche que nous proposons est illustrée dans la figure 3.2. Il s'agit d'une approche itérative dans laquelle chaque itération « t » de l'approche de raffinement proposée est constituée de trois étapes essentielles :

- L'estimation des distributions de possibilités à partir de l'ensemble $B = \bigcup_{k=0}^t B_k$ de zones d'apprentissage : cette estimation des distributions de possibilités est basée sur les transformations des distributions de probabilités vers des distributions de possibilités.
- La projection de l'image I sur ces distributions de possibilités permettant de générer des images appelées « cartes possibilistes » (CP_{I,C_m} , $m = 1, \dots, M$) qui représentent la possibilité d'appartenance de chaque pixel de l'image initiale I aux différentes classes thématiques de ψ .
- L'exploitation des cartes possibilistes ainsi que des informations spatiales par le biais d'un système de raisonnement possibiliste spatial afin de sélectionner de nouveaux germes possibilistes pour chacune des classes thématiques. Ces germes sont ajoutés à l'ensemble $B = \bigcup_{k=0}^{t+1} B_k$ pour l'estimation des distributions de possibilités à l'itération suivante « $t+1$ ».

Le processus d'enrichissement des zones d'apprentissage est ensuite répété itérativement jusqu'à ce qu'aucun germe ne puisse être rajouté à l'ensemble des zones d'apprentissage.

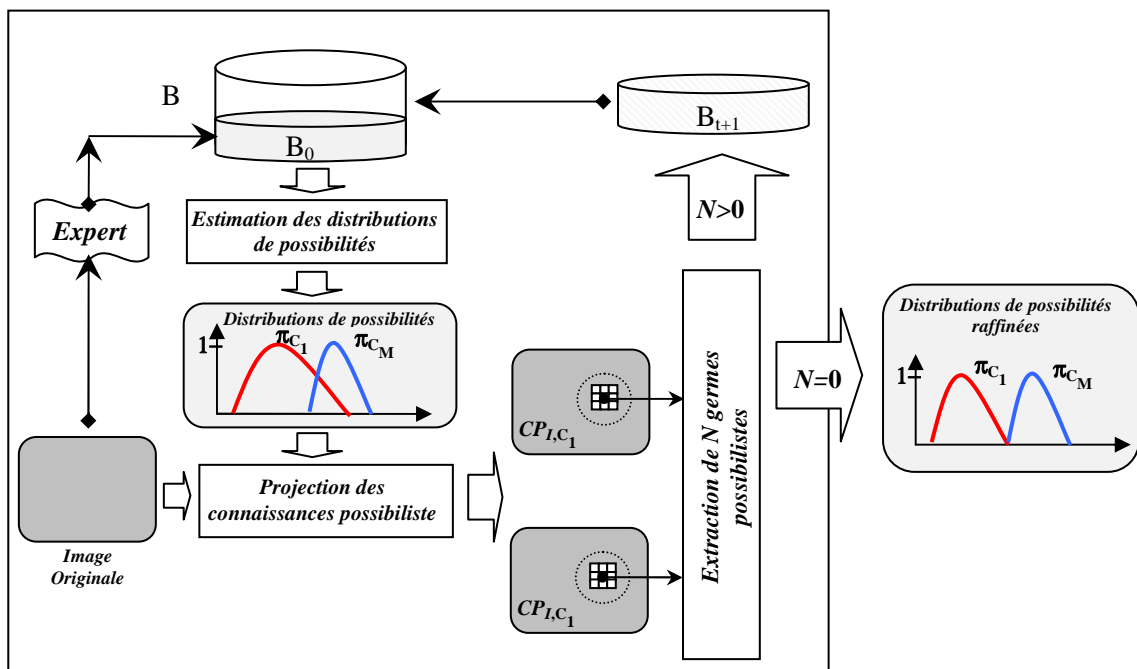


Figure 3.2 : Approche proposée pour le raffinement des distributions de possibilités

Dans la suite, nous allons présenter en détail, chacune des étapes de cette approche.

3.3.1 Estimation des distributions de possibilités

Comme nous l'avons déjà précisé, dans cette étude, nous considérons que la connaissance *a priori* est exprimée par l'expert à travers la définition de petites zones d'apprentissage considérées comme représentatives des différentes classes thématiques. L'objectif de ce sous-système est de réaliser, à partir de ces zones d'apprentissage, une estimation des distributions de possibilités aux différentes classes thématiques (Figure 3.3). La méthode que nous proposons d'appliquer comporte deux phases.

Dans un premier temps, la méthode d'estimation par noyau KDE (*Kernel Density Estimation*) [107], qui est une méthode non-paramétrique d'estimation de la densité de probabilité, est utilisée pour l'estimation des densités de probabilités des classes concernées à partir des zones d'apprentissages. En effet, la méthode KDE est particulièrement utilisée dans le cas où les données statistiques sont limitées. De plus, cette méthode, non-paramétrique, permet d'estimer la forme de la distribution de probabilités suivie par les données statistiques sans imposer d'allure à cette distribution. Bien souvent, le noyau utilisé dans cette méthode K est choisi comme la densité d'une fonction gaussienne standard (espérance nulle et variance unitaire) (Equation (3.8)). Généralement, le choix du noyau K est peu influent sur l'estimation des distributions de probabilités, mais il n'en est pas de même pour la largeur du noyau L (le paramètre de lissage). En effet, une valeur trop faible de L provoque l'apparition de détails artificiels apparaissant sur la distribution de probabilités estimée. Pour une valeur de L trop grande, la majorité des caractéristiques est au contraire effacée. Le choix de L est donc une question centrale dans l'estimation de distributions de probabilités. L'effet de ce choix sur la sélection des germes possibilistes sera étudié dans le paragraphe 3.4.3.

$$E(x) = \frac{1}{NL} \sum_{i=1}^N K\left(\frac{x-x_i}{L}\right) \quad (3.8)$$

Avec x une variable aléatoire prenant ses valeurs dans l'univers de discours $\Omega = \{x_1, x_2, \dots, x_N\}$ et $E(x)$ est la valeur estimée de la densité de probabilité à x .

Dans un deuxième temps, les densités de probabilités ainsi obtenues sont transformées en des distributions de possibilités en utilisant la transformation de probabilités-possibilités symétrique proposée par Dubois-Prade [90] [105][106].

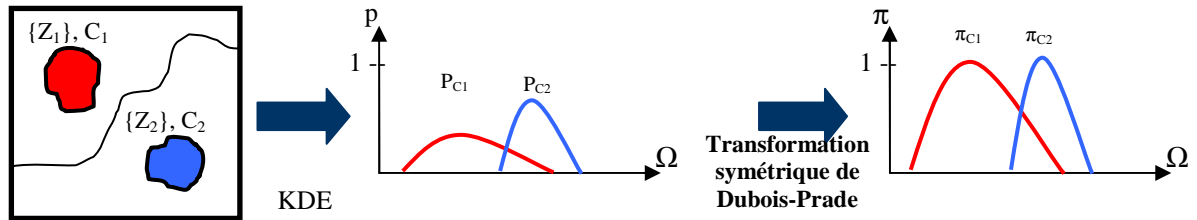


Figure 3.3 : Méthode d'estimation des distribution de possibilités à base de transformation probabilité-possibilité

3.3.2 Projection des connaissances possibilistes

L'étape d'estimation des distributions de possibilités aboutit à la construction de M distributions de possibilités (associées aux M classes thématiques $\psi = \{C_1, C_2, \dots, C_M\}$ présentes dans la scène analysée et prédéfinies par l'expert dans l'image observée I) :

$$\begin{aligned} \pi_{C_m} : \Omega &\rightarrow [0,1] \\ x &\rightarrow \pi_{C_m}(x) \end{aligned} \quad (3.9)$$

où x désigne la primitive observée.

$\pi_{C_m}(x(P))$ associe alors chaque pixel $P \in I$, ayant une valeur de primitive $x(P) \in \Omega$, à un degré de possibilité d'être de la classe C_m , $m = 1, \dots, M$. Etant donnée l'image I de taille $H \times W$, l'application de ces

distributions de possibilités sur cette image permet la projection des connaissances de l'expert, et ainsi, l'obtention de M images de taille H×W appelées cartes possibilistes CP_{I,C_m} , $m= 1, \dots, M$:

$$CP_{I,C_m}(i,j) = \pi_{C_m}(x(P(i,j))), i=1,\dots,H; j=1,\dots,W$$

Chaque carte possibiliste CP_{I,C_m} , $m= 1, \dots, M$ a les mêmes dimensions que l'image I , et comporte les degrés de possibilités pour tous les pixels de l'image I , d'être de la classe C_m .

Il est intéressant de remarquer que la projection des connaissances possibilistes constitue la première étape dans tous les systèmes de classification possibilistes d'images au niveau pixelique (Figure 3.4).

La deuxième étape dans de tels systèmes de classification consiste en l'application d'un raisonnement possibiliste et des mécanismes de décision sur les M cartes possibilistes afin de former l'image dite *thématique* où chaque pixel sera associé à une, et une seule, des classes thématiques prédéfinies par l'expert.

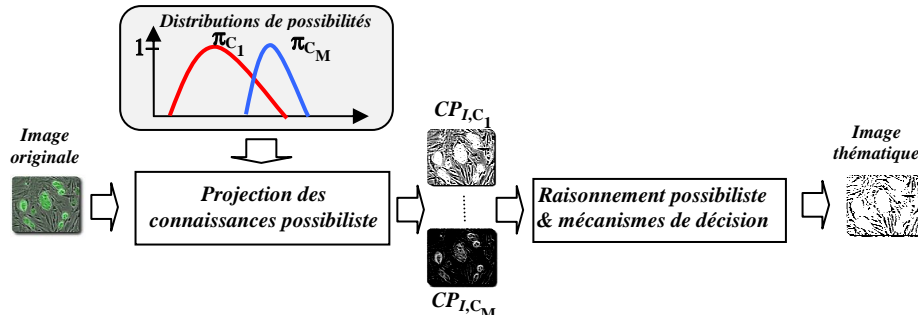


Figure 3.4 : Processus de classification pixelique des images

3.3.3 Sélection des germes possibilistes

Un aspect essentiel de l'approche proposée de raffinement des distributions de possibilités, réside dans l'intégration de nouvelles connaissances sous la forme de nouveaux échantillons dont la possibilité d'appartenance aux différentes classes thématiques est certaine. Ces échantillons sont appelés les *germes possibilistes*. Ce processus est fortement inspiré des approches de segmentation d'images par croissance de régions (i.e. sélection et « intégration » de pixels ayant des valeurs spectrales similaires à des zones représentatives voisines).

Un germe possibiliste associé à une classe donnée C_m est défini comme étant un pixel pour lequel l'état des connaissances possibilistes disponibles permet d'affirmer avec la plus grande certitude que ce pixel appartient à la classe C_m .

Dans ce travail, nous proposons une démarche fortement inspirée de la vision humaine pour la sélection des germes. En effet, l'analyse humaine concernant la possibilité d'appartenance aux différentes classes d'un pixel se passe en deux temps.

Dans un premier temps, et relativement à chaque classe thématique, le contexte spatial « local » est considéré comme une source de connaissance permettant d'ajuster le degré de possibilité pour un pixel concerné d'appartenir à chaque classe thématique. Ce processus, dit de *focalisation d'attention spatiale*, consiste donc, pour chaque pixel considéré, à estimer un degré de possibilité « spatial » conditionnellement aux différentes classes thématiques.

La deuxième phase, dite *d'appréciation décisionnelle spatiale*, consiste à apprécier la certitude, ou le niveau de confiance, d'une prise de décision pour l'attribution d'une étiquette de classe à chaque pixel en confrontant les nouveaux degrés de possibilité conditionnellement aux autres classes thématiques.

En considérant l'ensemble des classes thématiques $\psi = \{C_1, C_2, \dots, C_M\}$, les cartes possibilistes CP_{I,C_m} , $m=1, \dots, M$, et pour chaque pixel $P_0 \in I$, la mise en place de cette démarche est réalisée de la manière suivante :

- **Phase de focalisation d'attention spatiale :**

La distribution spatiale de possibilités $\widetilde{\pi}_{P_0} = [\widetilde{\pi}_{P_0}(C_1), \widetilde{\pi}_{P_0}(C_2), \dots, \widetilde{\pi}_{P_0}(C_M)]$ (définie sur l'ensemble des classes thématiques ψ) est obtenue par l'application d'un opérateur Φ (l'opérateur de moyenne, par exemple) sur le voisinage spatial $V(P_0)$ centré sur le pixel P_0 au niveau des différentes cartes possibilistes CP_{I,C_m} , $m=1, \dots, M$:

$$\widetilde{\pi}_{P_0}(C_m) = \Phi_{P \in V(P_0)}(CP_{I,C_m}(P)) \quad (3.10)$$

- **Phase d'appréciation décisionnelle spatiale :**

Sur la base de la distribution spatiale de possibilités, un indice de confiance spatial est établi pour chaque classe C_m :

$$\widetilde{Ind}(A_m) = \widetilde{\pi}_{P_0}(C_m) - \max_{n \neq m} [\widetilde{\pi}_{P_0}(C_n)] \quad (3.11)$$

où $A_m = \{C_m\}$, $m=1, \dots, M$. Rappelons à ce stade que cet indice de confiance permet de quantifier l'incertitude liée à la réalisation de chaque événement élémentaire A_m , $m=1, \dots, M$ (i.e. l'incertitude liée à l'appartenance du pixel P_0 à chacune des classes thématiques).

Dans le cas où $\Phi = I_{P_0}$ (i.e. l'opérateur d'identité et $V(P_0) = \{P_0\}$), la distribution spatiale de possibilités est tout simplement réduite à la distribution de possibilités pixelique représentée dans les cartes possibilistes : $\pi_{P_0} = [\pi_{P_0}(C_1), \pi_{P_0}(C_2), \dots, \pi_{P_0}(C_M)]$, où $\widetilde{\pi}_{P_0}(C_m) = \pi_{P_0}(C_m) = CP_{I,C_m}(P_0)$.

Dans ce cas, l'indice spatial se réduit à un indice *Ind* dit pixelique.

La règle de prise d'une décision que nous proposons alors, pour qu'un pixel considéré soit un germe possibiliste ou non, est basée sur les deux niveaux pixelique et spatial à la fois. En effet, en considérant l'indice de confiance pixelique *Ind* (établi à partir de la distribution de possibilités pixelique $\pi_{P_0} = [\pi_{P_0}(C_1), \pi_{P_0}(C_2), \dots, \pi_{P_0}(C_M)]$), la règle de décision proposée est la suivante :

$$P_0 \in I \text{ est un germe possibiliste ssi :} \\ \exists C_{m_0} \in \psi / \begin{cases} \widetilde{Ind}(\{C_{m_0}\}) \geq S \\ \widetilde{Ind}(\{C_{m_0}\}) = \widetilde{\pi}_{P_0}(C_{m_0}) - \max_{m \neq m_0} \widetilde{\pi}_{P_0}(C_m) \geq S \end{cases} \quad (3.12)$$

où S est un seuil de confiance décisionnelle appliqué sur les deux niveaux pixelique et spatial. Notons que, l'ensemble des germes possibilistes sélectionnés en exploitant conjointement les deux niveaux pixelique et spatial est en réalité un sous-ensemble des germes possibilistes

sélectionnés au niveau pixelique seul pour lesquels, l'utilisation conjointe des indices de confiance pixelique et spatiale est effectuée dans le but d'augmenter la certitude dans le processus d'extraction des germes possibilistes.

Finalement, l'approche de raffinement des distributions de possibilités proposée est résumée de la façon algorithmique suivante :

Entrée :

- La sélection par un expert des zones d'apprentissage $B_0 = \{Z_{0,1}, Z_{0,2}, \dots, Z_{0,M}\}$;
- La base d'échantillons d'apprentissage est initialisée par les zones d'apprentissage ;
- Le critère d'arrêt des itérations est mis à la position de « non vérifié ».

Tant que (Le critère d'arrêt des itérations est non vérifié) **Faire**

➤ **Première étape : Calcul des cartes possibilistes**

- Estimation des distributions de probabilités pour chacune des classes thématiques prédéfinies dans la scène par l'application de la méthode du noyau KDE sur la base d'apprentissage ;
- Transformation des distributions de probabilités en des distributions de possibilités par l'application de la transformation symétrique proposée par Dubois-Prade ;
- Calcul des M cartes possibilistes correspondant aux différentes classes thématiques : $CP_{l,C_m}(i,j) = \pi_{C_m}(x(P(i,j)))$, $i=1, \dots, H$; $j=1, \dots, W$.

➤ **Deuxième étape : Détermination des germes possibilistes**

• **Pour** chaque pixel P de l'image **Faire**

- Extraction d'un voisinage $V(P)$ (par exemple 3×3 pixels) centré sur P pour l'ensemble des cartes possibilistes ;
- Calcul de la distribution spatiale de possibilités par l'application de l'opérateur Φ (moyenne) sur l'ensemble des cartes possibilistes ;
- Calcul des indices de confiance pixelique Ind et spatial \widetilde{Ind} pour l'ensemble des classes thématiques ;
 - **Si** (Il existe une classe thématique C_{m0} pour laquelle $\widetilde{Ind} > S$ et $Ind > S$) **Faire**
 - Ajouter le pixel P comme germe possibiliste à l'ensemble de germes $Z_{m0} = \bigcup_{k=0}^{t+1} Z_{k,m0}$ de la classe C_{m0}
 - **Fin-Si**

• **Fin-Pour**

➤ **Troisième étape : Vérification du critère d'arrêt**

- Critère d'arrêt adopté : L'évolution du nombre de germes sélectionnés par chaque itération. Si ce nombre n'est pas nul, une nouvelle itération est lancée, sinon c'est la fin de la boucle des itérations.

Fin - Tant que

Sortie : Distributions de possibilités raffinées.

3.4 Expérimentations et résultats

Dans cette section, l'évaluation de l'efficacité de l'approche itérative proposée en termes de caractérisation des classes prédéfinies dans une scène analysée, est étudiée. Deux cas de discrimination de classes sont principalement considérés. Il s'agit des cas de faible et de grande complexité de discrimination des classes en termes de classification pixelique. Dans la section 3.4.1, un ensemble d'images de synthèse est généré à des fins d'évaluation de convergence de l'approche itérative proposée. Les sections suivantes abordent l'efficacité de la méthode de sélection des germes possibilistes et l'impact du seuil sur le processus de raffinement.

3.4.1 Présentation des données expérimentales

Dans le but d'évaluer l'approche proposée, une image de synthèse de taille 96×128 pixels, illustrée dans la Figure 3.5-(a), a été générée. Cette image de synthèse est composée de deux classes thématiques $\{C_1, C_2\}$. La classe C_1 présente dans la scène sous la forme de quatre objets d'intérêt, circulaires de tailles différentes; la classe C_2 constitue le fond de l'image. Nous faisons l'hypothèse que les classes d'objets présents sur l'image analysée suivent des distributions de probabilités gaussiennes. Les pixels sont donc générés à partir de variables aléatoires gaussiennes $G(m_1, \sigma_1)$, pour la classe de C_1 , et $G(m_2, \sigma_2)$ pour la classe C_2 .

Afin de générer une image représentant le cas d'une faible complexité de discrimination de classes, les paramètres suivants sont utilisés (Figure 3.5-(b)) : $G(160, 10)$ pour la classe C_1 , et $G(100, 15)$ pour la classe C_2 . La notion de faible complexité de discrimination est interprétée (dans notre travail) par un écart important des valeurs moyennes et des petites valeurs d'écart-type des distributions de probabilités. D'une façon similaire, l'utilisation des paramètres suivants $G(120, 15)$ pour la classe C_1 et $G(100, 20)$ pour la classe C_2 permet de représenter le cas de grande complexité de discrimination entre les classes (Figure 3.5-(c)). La notion de grande complexité de discrimination est interprétée (dans notre travail) par un faible écart des valeurs moyennes des distributions de probabilités et une augmentation des valeurs d'écart type. Ceci implique donc un faible contraste entre les classes à discriminer ce qui rend complexe la classification au niveau pixelique.

Des zones d'apprentissage de taille réduite (10×10 pixels) sont positionnées par l'expert (comme étant des zones représentatives des classes thématiques considérées) et indiquées sur l'image générée. En utilisant ces zones initiales d'apprentissage, l'estimation initiale des distributions de probabilités des différentes classes thématiques est réalisée par la méthode d'estimation KDE. L'application de la transformation $Pr-\pi$ symétrique de Dubois-Prade permet d'obtenir les distributions de possibilités initiales qui sont également représentées dans la Figure 3.5.

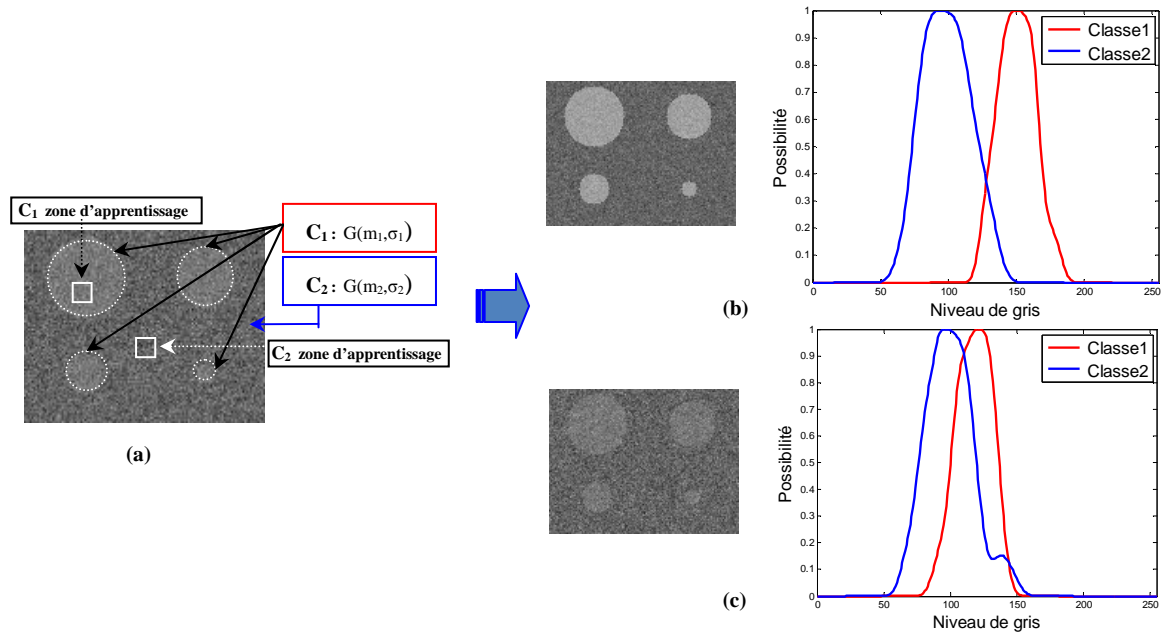


Figure 3.5 : (a) Images de synthèse avec les zones d'apprentissage (b) image de synthèse (cas de faible complexité de discrimination entre classes) et les distributions de possibilités initiales, (c) image de synthèse (cas de grande complexité de discrimination entre classes.) et les distributions de possibilités initiales

Comme nous pouvons le constater, dans le cas de faible complexité de discrimination des classes, l'allure gaussienne (aussi appelée en cloche) des distributions de possibilités obtenues est bien préservée, tandis que dans le cas de grande complexité de discrimination de classes et à cause du fort niveau de chevauchement, l'allure gaussienne des distributions de possibilités n'est pas totalement préservée. Notons que, dans la Figure 3.5-(c), le soubresaut pour la classe C_2 est dû au nombre réduit d'échantillons engendrés d'une façon aléatoire et surtout à l'estimation par la méthode KDE (i.e. problème du choix de la largeur du noyau L).

3.4.2 Evaluation de la méthode de sélection des germes possibilistes

Dans cette section, la méthode proposée de sélection des germes possibilistes est étudiée en fonction du seuil de confiance S . Une fenêtre de taille 3×3 pixels centrée sur chaque pixel est considérée comme le contexte spatial de ce pixel. Une fois les zones d'apprentissage initiales délimitées par l'expert, l'approche itérative proposée est appliquée à l'image de synthèse dans les deux cas considérés (i.e. cas de faible et de grande complexité de discrimination entre les classes). Dans la Figure 3.6, le nombre de germes correctement sélectionnés ainsi que de germes erronément sélectionnés est donné pour les deux cas considérés, en fonction du seuil $S \in [0,1]$ (notons que c'est grâce à l'utilisation d'images synthétiques que nous pouvons apprécier le choix correct ou non, des germes sélectionnés).

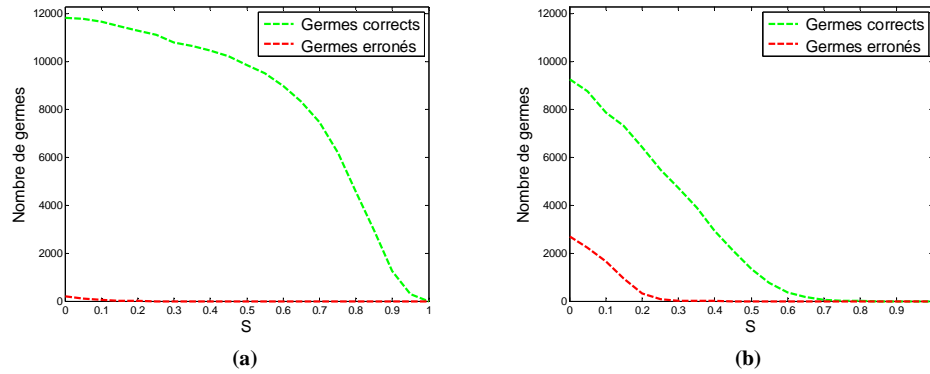


Figure 3.6 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil dans le cas gaussien: (a) cas de faible complexité de discrimination entre les classes (b) cas de grande complexité de discrimination entre les classes

Dans le contexte de faible complexité de discrimination entre les classes, la règle de sélection des germes possibilistes (Equation (3.12)) permet d'extraire un nombre important de germes corrects et un faible nombre de germes erronés et ceci même pour des faibles valeurs du seuil S (Figure 3.6-(a)).

En revanche, dans le cas d'un contexte de grande complexité de discrimination entre les classes, le nombre des germes corrects sélectionnés diminue rapidement en fonction de seuil de confiance S et le nombre de germes erronés est augmenté d'une façon importante pour les faibles valeurs du seuil S (Figure 3.6-(b)).

Notons que notre objectif est d'extraire avec la plus grande certitude, des germes corrects. Par conséquent, le choix d'une valeur du seuil permettant de réduire considérablement l'extraction des germes possibiliste « erronés » est d'une importance primordiale. Pour montrer l'effet du seuil sur les distributions de possibilités raffinées, nous avons appliqué l'approche proposée sur l'image de synthèse dans le contexte de grande complexité de discrimination entre les classes et ceci pour les valeurs suivantes de seuil : $S = 0.1, 0.5$, et 0.8 . Les résultats ainsi obtenus sont donnés dans la Figure 3.7.

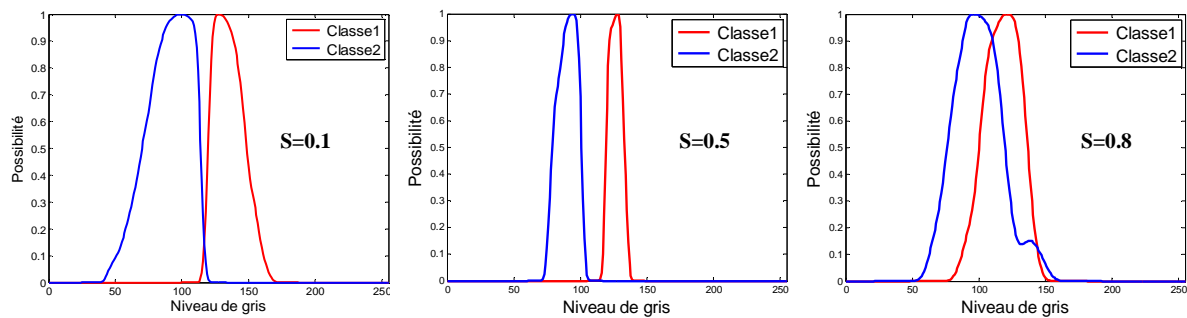


Figure 3.7 : Effet de la valeur de seuil S sur les distributions de possibilités raffinées dans le cas de grande complexité de discrimination entre classes et pour $S=0.1, 0.5$ et 0.8

Comme nous pouvons le remarquer, le choix d'une faible valeur de seuil ($S=0.1$) implique l'intégration d'un grand nombre de germes erronés. En revanche, le choix d'une grande valeur de seuil de confiance (le cas de $S=0.8$, illustré dans la figure 3.7) ne permet de sélectionner qu'un nombre extrêmement faible et non significatif de germes. Par conséquent, les distributions de possibilités raffinées sont pratiquement inchangées et se confondent avec les distributions de possibilités initiales (estimées à partir des zones d'apprentissages fixées par l'expert).

En conclusion, dans les deux cas de contextes de faible et de grande complexité de discrimination entre les classes, la sélection d'une valeur du seuil S supérieur à 0.3 produit un nombre extrêmement

faibles de germes erronés tout en aboutissant à un nombre raisonnable de germes corrects dans le cas de contexte de grande complexité. Pour la suite, nous adoptons le choix d'une valeur de seuil $S=0.5$ qui semble aboutir à un compromis raisonnable entre le nombre de germes corrects (avec une grande confiance), et un très faible nombre de germes erronés. La Figure 3.8 montre le positionnement spatial des germes possibilistes sélectionnés à l'issue du processus de raffinement et ceci pour les deux classes thématiques dans le cas de grande complexité de discrimination de classes et pour une valeur de seuil $S=0.5$. Nous pouvons constater que les germes possibilistes de la classe C_1 (resp. de la classe C_2) sont positionnés dans les zones de la classe C_1 (resp. de la classe C_2), en d'autres termes, pas de germes erronés pour les deux classes C_1 et C_2 .

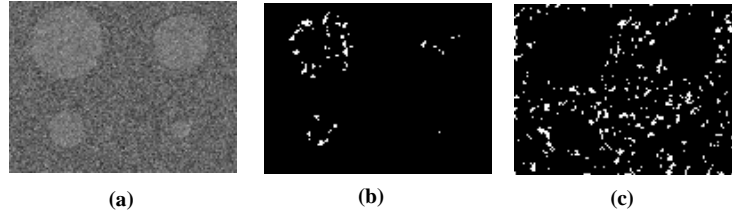


Figure 3.8 : (a)Image de synthèse (cas de grande complexité de discrimination entre classes.), (b) germes possibilistes sélectionnés pour la classe C_1 , (c) germes possibilistes sélectionnés pour la classe C_2

3.4.3 Effet de la largeur L du noyau de la méthode KDE

La largeur du noyau L est un paramètre qui présente une forte influence sur le résultat de l'estimation de distributions de probabilités. Pour les distributions de probabilités gaussiennes, le choix optimal de ce paramètre est donné par [108]:

$$L = \left(\frac{4\sigma^5}{3n} \right)^{\frac{1}{5}} \quad (3.13)$$

où σ est l'écart type des n échantillons. Ce choix optimal a été utilisé dans l'évaluation de la méthode de sélection des germes possibilistes dans le cas des classes ayant des distributions de probabilités gaussiennes (sous-section 3.4.2). Afin de montrer l'effet de ce paramètre dans l'évaluation de la méthode de sélection des germes possibilistes (dans le cas des distributions de probabilités de Rayleigh ainsi que le cas des distributions multimodales gaussiennes), l'image de synthèse illustrée dans la Figure 3.5-(a) a été utilisée. Dans un premier temps, les pixels sont générés à partir de variables aléatoires de Rayleigh $R(\alpha_1)$, pour la classe C_1 , et $R(\alpha_2)$ pour la classe C_2 et, dans un deuxième temps, à partir de variables aléatoires multimodales gaussiennes $G(m_1, \sigma_1)$ et $G(m_2, \sigma_2)$ pour la classe de C_1 . Dans les deux cas Rayleigh et multimodales, les paramètres utilisés sont les suivants :

- Le cas de Rayleigh de faible complexité de discrimination entre les classes (avec les paramètres suivants $\{\alpha_1(100), \alpha_2(30)\}$), et le cas Rayleigh de grande complexité de discrimination entre les classes (avec les paramètres suivants $\{\alpha_1(65), \alpha_2(30)\}$).
- Le cas de multimodales gaussien de faible complexité de discrimination entre les classes (avec les paramètres suivants $\{G(50, 10)$ et $G(150, 10)\}$), et le cas multimodales gaussien de grande complexité de discrimination entre les classes (avec les paramètres suivants $\{G(100, 10)$, resp. $G(100, 15)\}$).

Quatre valeurs de la largeur du noyau 0.5, 2, 7 et la valeur optimale (i.e. calculé à partir de l'Equation (3.13)) ont été testées dans la méthode de sélection des germes possibilistes et les résultats sont montrés dans : Figure 3.9, Figure 3.10, Figure 3.11 et Figure 3.12

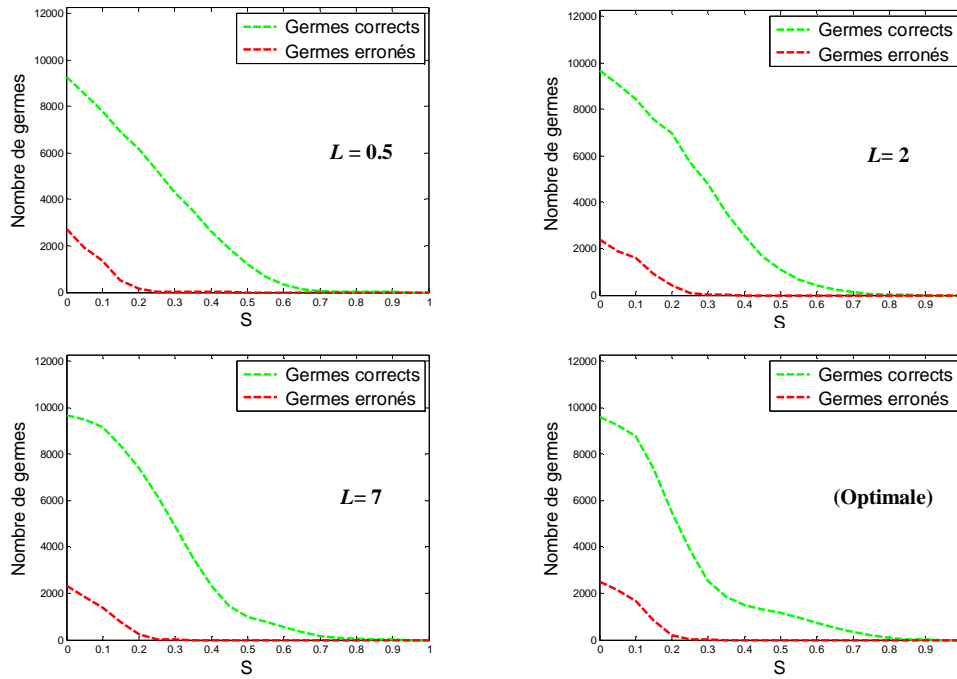


Figure 3.9 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de Rayleigh avec une faible complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale

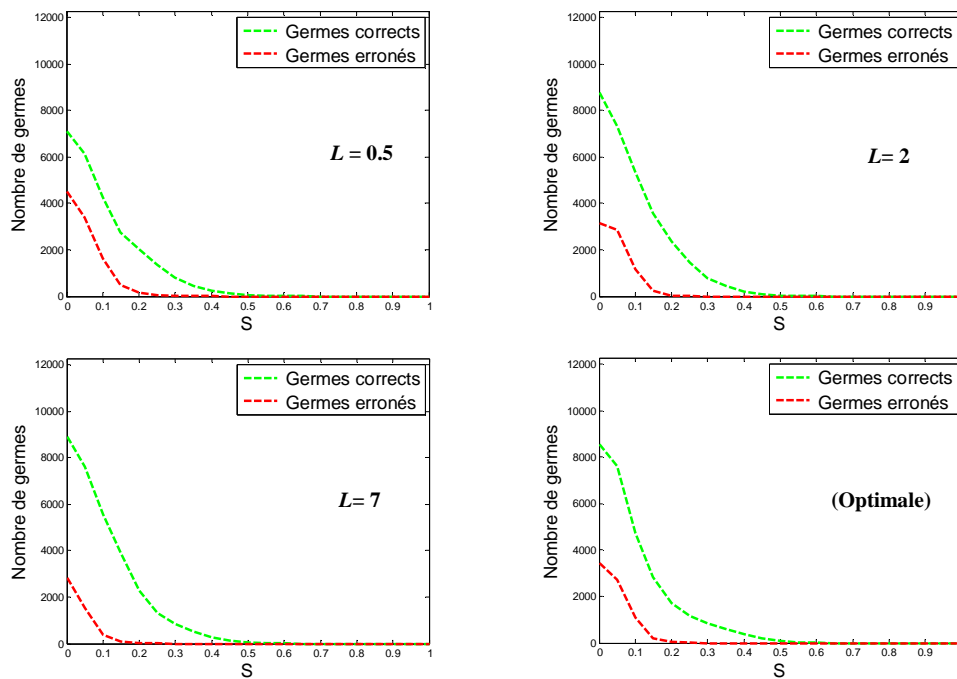


Figure 3.10 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de Rayleigh avec une grande complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale

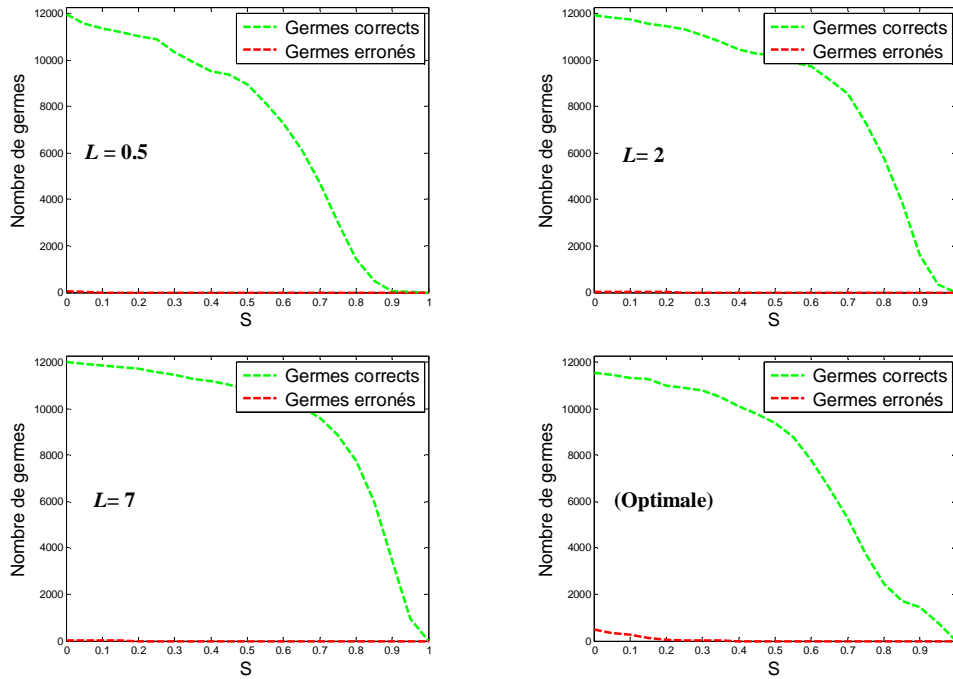


Figure 3.11 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de multimodales gaussien avec une faible complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale

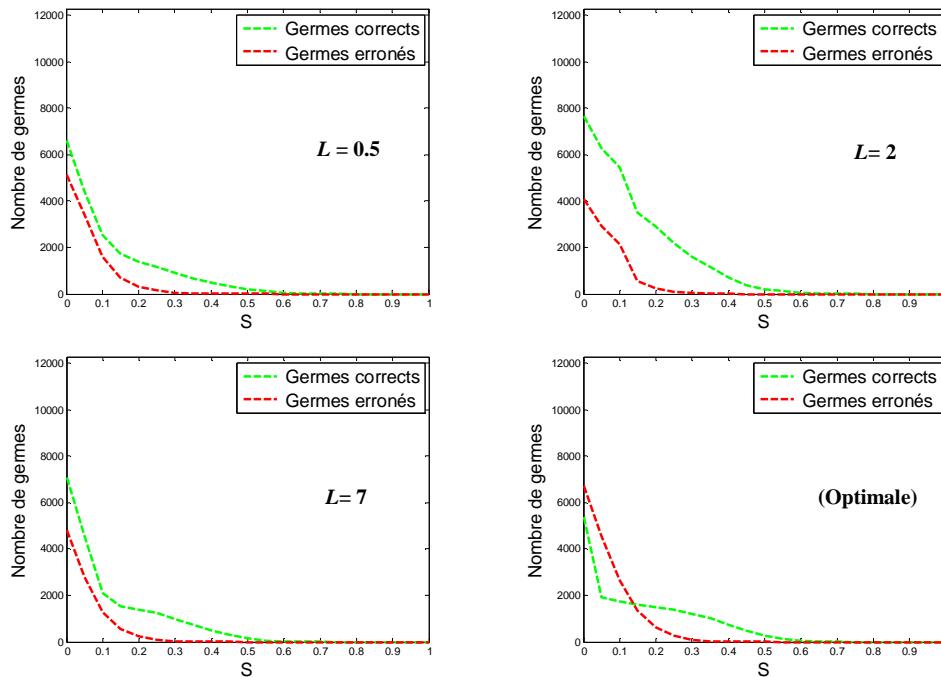


Figure 3.12 : Nombre de germes (corrects et erronés) sélectionnés par la méthode d'extraction de germes en fonction du seuil, dans le cas de multimodales gaussien avec une grande complexité de discrimination entre les classes et pour quatre valeurs de la largeur du noyau 0.5, 2, 7 et valeur optimale

On constate que dans le cas de Rayleigh (avec faible et grande complexité de discrimination entre les classes et pour une valeur de la largeur $L=2$) que la valeur de seuil $S=0.3$ permet de sélectionner le nombre de germes corrects (avec une grande confiance) le plus élevé et sans avoir de germes erronés. Dans le cas de multimodales gaussiens (avec faible et grande complexité de discrimination entre les classes et pour une valeur de la largeur $L=2$), la même valeur de seuil $S=0.3$ peut être adoptée afin d'avoir un compromis raisonnable entre le nombre de germes corrects et le faible nombre de germes

erronés. Ces valeurs de largeur du noyau $L=2$ ainsi que de seuil $S=0.3$ seront adoptées dans la suite de ce travail pour traiter le cas de Rayleigh et le cas de multimodales gaussiens.

3.4.4 Comportement de l'approche proposée

Dans cette section, la qualité des distributions de possibilités raffinées est évaluée dans les deux contextes de faible et de grande complexité de discrimination entre les classes thématiques présentes dans la scène analysée. Pour pouvoir mener cette évaluation, nous avons besoin d'avoir des distributions de possibilités de référence permettant d'apprécier la qualité des distributions de possibilités obtenues après raffinement.

Etant donné que nous utilisons des images de synthèse pour lesquelles les distributions de probabilités sont parfaitement connues *a priori*, nous proposons de considérer comme référence, les distributions de possibilités obtenues par l'application de la méthode de transformation symétrique de Dubois et Prade sur les distributions de probabilités $G(m_1, \sigma_1)$ et $G(m_2, \sigma_2)$.

Dans la Figure 3.13, les distributions de possibilités initiale et après raffinement sont données et comparées par rapport aux distributions de référence dans les deux contextes de faible et de grande complexité de discriminations de classes.

a. Informations extraites par cette approche

Une analyse des résultats obtenus montre deux situations différentes :

- Lorsque l'approche proposée est appliquée pour l'estimation des distributions de possibilités dans le cas où les classes thématiques sont relativement « faciles » à discriminer, alors les distributions de possibilités raffinées convergent vers les distributions de possibilités de référence (Figure 3.13- (c)).
- Lorsque les classes thématiques sont difficiles à discriminer (situation rencontrée lorsque les deux classes ont un fort niveau de chevauchement entre leurs distributions de possibilités), nous remarquons une convergence « sélective » des distributions de possibilités se focalisant sur l'extraction des germes les plus discriminants (c'est-à-dire, les germes les plus « fréquents » avec une grande confiance). Le fait que ces germes se trouvent vers les centres des distributions implique une déviation des distributions de possibilités raffinées afin de s'éloigner des zones de chevauchement entre les deux classes.

Néanmoins, il est important de noter que dans les deux cas, les distributions de possibilités obtenues (après raffinement) tronquent les « queues des distributions » (dus à la forme gaussienne des distributions initiales) et se focalisent sur les germes discriminants les plus « fréquents » dans la scène. En d'autres termes, le processus proposé d'extraction des germes possibilistes (qui s'inspire de la focalisation d'attention de la vision humaine) n'accorde pas d'importance aux germes en queue de distributions même si ceux-ci font une bonne discrimination entre les différentes classes.

Ces résultats soulèvent la question importante de la signification physique des distributions de possibilités utilisées pour la classification pixelique d'images et construites grâce à l'approche itérative proposée. En effet, dans le cas d'une faible complexité de discrimination entre les classes, les distributions de possibilités raffinées encapsulent une signification de *caractérisation des différentes classes* ; tandis que dans le cas des classes relativement difficiles à discriminer, les distributions de possibilités raffinées sont construites de sorte à réaliser une meilleure discrimination (et non une

meilleure caractérisation) des classes considérées. Cette interprétation explique la déviation des formes des distributions de possibilités raffinées par rapport aux distributions de possibilités de référence.

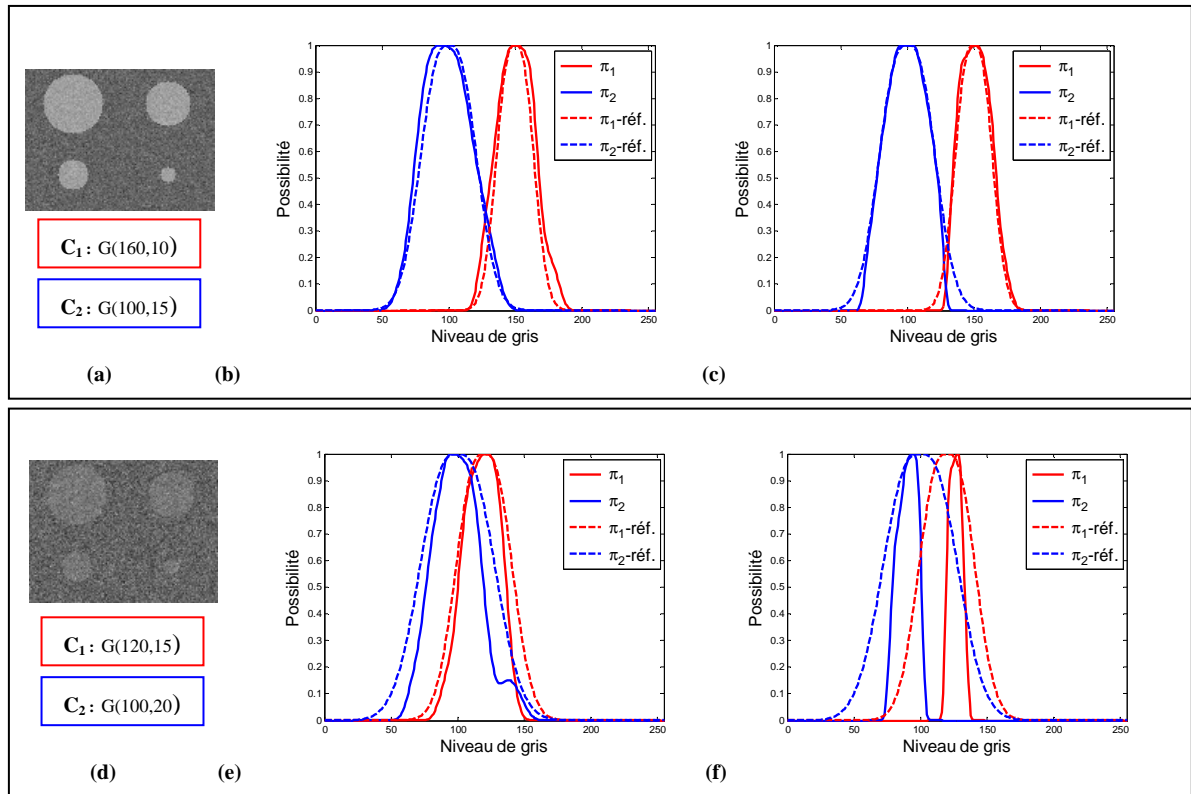


Figure 3.13 : (a) Image de synthèse (cas de faible complexité de discrimination), (b) les distributions de possibilités référence, (c) les distributions de possibilités référence et raffinée, (d) Image de synthèse (cas de grande complexité de discrimination), (e) les distributions de possibilités référence, (f) les distributions de possibilités référence et raffinée

b. Convergence de l'approche

Sur le plan de l'évaluation quantitative, la Figure 3.14 représente l'évolution du nombre des germes sélectionnés dans les deux cas considérés, en fonction du nombre d'itérations de raffinement des distributions de possibilités.

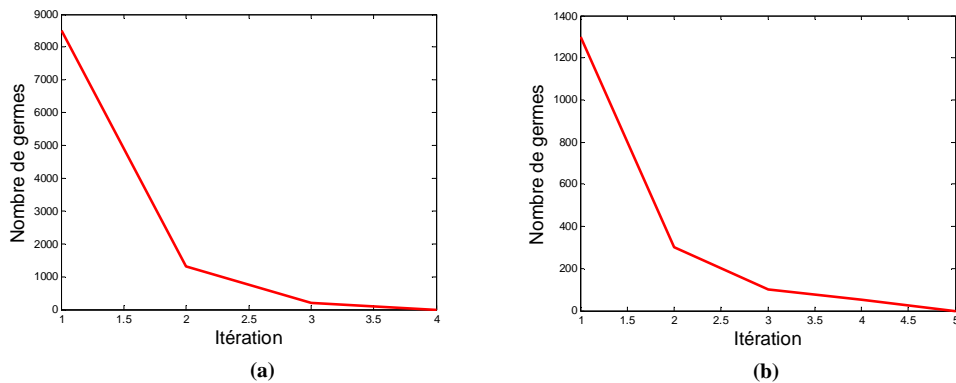


Figure 3.14 : (a) Nombre de germes sélectionnés en cas de faible complexité de discrimination entre les classes et (b) nombre de germes sélectionnés en cas de grande complexité de discrimination entre les classes

Etant donné que le volume des germes sélectionnés est utilisé comme critère d'arrêt du processus de raffinement, les résultats obtenus permettent de constater la convergence rapide, en moins de cinq itérations, de l'approche proposée et ceci dans les deux contextes de faible et de grande complexité de discrimination entre les classes.

3.4.5 Résultats expérimentaux sur des images médicales

Dans ce paragraphe, les performances de l'approche proposée en termes de caractérisation et discrimination entre les classes dans des images médicales est évaluée à titre purement illustratif. A cet effet, deux images sont seulement utilisées : une image d'une tumeur du sein obtenue par imagerie IRM « *imagerie par résonance magnétique* » (Breast MRI dataset RIDER (<http://www.cancerimagingarchive.net/>)), et une image mammographique extraite de la base d'images MIAS fournie par la société savante anglaise (*Mammographic Image Analysis Society*) (<http://www.mammoimage.org/databases/>) (Figure 3.11). Chaque image comporte deux classes : tissu normal et tissu tumoral. Dans le cas de la première image (Test-1), les deux classes peuvent être facilement discriminées et dans le deuxième cas (Test-2), il s'agit d'une image dans laquelle les classes sont difficiles à discriminer.

L'approche proposée de raffinement des distributions de possibilités a été appliquée sur les deux images Test-1, 2 où deux zones d'apprentissage de taille 10×10 pixels sont positionnées par un expert, sur chacune de ces images (comme étant des zones représentatives des deux classes: tissu normal et tissu tumoral). Ces zones d'apprentissage ainsi que les deux distributions de possibilités initiales et raffinées sont illustrées dans la Figure 3.16.

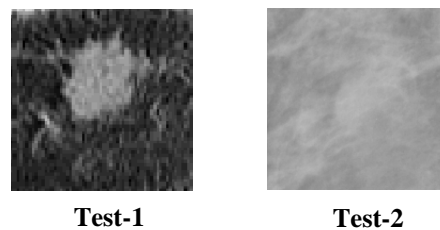


Figure 3.15 : (Test-1) Image d'une tumeur du sein obtenue par imagerie IRM et (Test-2) une image mammographique

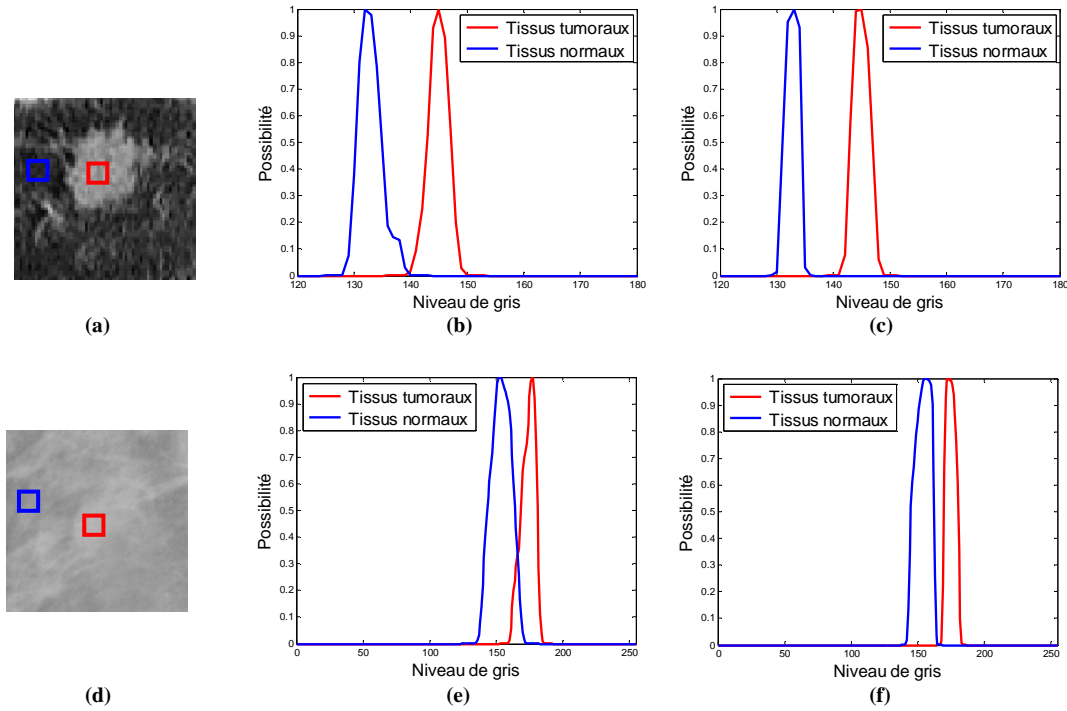


Figure 3.16 : Effet de raffinement des distributions de possibilités
 (a) et (d) image IRM et image mammographique respectivement avec deux zones d'apprentissage
 (b) et (e) distributions de possibilité initiales
 (c) et (f) distributions de possibilités raffinées

Comme nous pouvons le constater, l'effet de raffinement est moins remarquable dans le cas facile à discriminer entre les classes tandis que cet effet devient plus « visible » dans le cas de grande difficulté de discrimination entre les classes (où le processus de raffinement permet de réduire l'ambiguïté, ou la confusion, interclasses).

L'impact de ce processus de raffinement des distributions de possibilités sera abordé plus en détail dans le chapitre 5 qui abordera la question de segmentation d'image.

3.5 Conclusions

Dans ce chapitre, nous avons présenté une brève synthèse des méthodes d'estimation des distributions de possibilités en les répartissant en trois catégories : les méthodes d'estimation guidées par l'expert, les méthodes de transformation d'une distribution de probabilités en une distribution de possibilités et les méthodes basées sur de l'apprentissage. Nous avons distingué, dans la première catégorie, les approches heuristiques, des approches hybrides. Ces dernières d'une part intègrent l'opinion de l'expert pour la représentation des connaissances ambiguës en sélectionnant les formes standards des distributions de possibilités, et d'autre part, permettent la détermination des paramètres de ces distributions par l'intermédiaire d'une optimisation de mesures floues.

Un intérêt particulier a été porté aux méthodes de transformation d'une distribution de probabilités en une distribution de possibilités. Cette deuxième catégorie est très intéressante car elle permet l'intégration des données statistiques. Nous avons abordé le principe général de ces méthodes en se focalisant sur la méthode proposée par Dubois-Prade permettant la transformation symétrique. Cette transformation a été adoptée dans notre travail de recherche à cause des bonnes performances de cette méthode de transformation dans le domaine de la reconnaissance de formes.

D'une manière générale, la majorité des méthodes d'estimation des distributions de possibilités tentent de caractériser les classes à partir d'un ensemble limité d'échantillons (données statistiques extraites à partir de zones représentatives, délimitées par un expert sur une image). Ceci constitue un inconvénient majeur de ces méthodes surtout dans un contexte de confusion interclasses (grande complexité de discrimination entre les classes).

Une approche de raffinement des distributions de possibilités a été proposée dans ce chapitre. Cette approche procède en trois étapes : une estimation des distributions de possibilités à partir d'un ensemble limité de données, une projection de ces distributions de possibilités sur l'image étudiée et enfin l'application d'un mécanisme de raisonnement possibiliste exploitant le contexte spatial et permettant de sélectionner des nouveaux germes possibilistes (qui seront intégrés dans l'ensemble des échantillons d'apprentissage).

Dans la première étape, une approche d'estimation des distributions de possibilités a été adoptée. Cette approche est basée sur l'application de la technique KDE afin d'estimer une distribution de probabilités, et dans un deuxième temps, en l'application de la transformation de Dubois Prade symétrique permettant d'obtenir les distributions de possibilités pour les différentes classes thématiques présentes dans la scène analysée. Ces distributions de possibilités sont ensuite utilisées pour réaliser une projection possibiliste (liée aux différentes classes thématiques) des connaissances contenues dans l'image analysée. Un ensemble de cartes possibilistes est obtenu à l'issue de cette projection.

L'application d'un mécanisme de raisonnement possibiliste est ensuite effectuée sur les cartes possibilistes dans l'objectif d'enrichir l'ensemble des données disponibles par la sélection de nouveaux germes appartenant avec une grande certitude aux différentes classes considérées dans la scène. Ces nouveaux germes sont injectés dans l'ensemble d'apprentissage d'une façon itérative permettant ainsi de réaliser un raffinement des distributions de possibilités estimées.

L'approche proposée pour le raffinement des distributions de possibilités a été appliquée aux images de synthèse ainsi qu'à des images médicales (IRM et mammographie) comportent deux classes. Les résultats obtenus montrent que l'ambiguïté entre les classes a été réduite considérablement. L'impact de cette réduction de l'ambiguïté sur la segmentation d'images sera approfondi dans le chapitre 5.

Chapitre 4

SIMILARITE POSSIBILISTE

4.1 Introduction

La notion de similarité représente un concept essentiel et un outil fondamental dans le raisonnement humain face aux problèmes de la vie quotidienne comme la compréhension du langage, la reconnaissance des formes ou la prise de décision. Les humains forment un modèle mental de certains concepts à travers leurs expériences. Ils utilisent ce modèle comme la base pour un raisonnement analogique qui consiste en la mise en correspondance d'une situation (ou d'un objet) antérieure et d'une situation nouvelle qui lui est « similaire », afin d'en déduire une décision, une interprétation, etc. correspondant à la situation en cours [118].

David Hume [120] a exprimé ce type de comportement par le principe d'extrapolation : « *In reality, all arguments from experience are founded on the similarity, which we discover among natural objects, and by which we are induced to expect effects similar to those, which we have found to follow from such objects. From causes, which appear similar, we expect similar effects. This is the sum of all our experimental conclusions* ». Ainsi, Hume suggère d'extrapoler certaines propriétés d'un objet parfaitement bien connu vers d'autres objets qui lui sont similaires.

Le concept de similarité a fait l'objet d'importants travaux de recherche dans les différents domaines du traitement de l'information tels que la reconnaissance de formes, la fouille des données et l'intelligence artificielle [121][122] (par exemple, pour le raisonnement à base de cas [119], le raisonnement analogique [118], etc.). Rissland [122] a souligné l'importance et le rôle central du concept de similarité en l'expliquant par la difficulté de représenter des concepts du monde réel. Pour résumer, l'utilisation de la similarité a été menée dans le but de transcrire le comportement humain et motivée par l'idée qu'il est généralement très difficile de généraliser à partir d'un ensemble d'apprentissage limité.

Plusieurs méthodes de reconnaissance de formes sont fortement basées sur l'utilisation du concept de similarité. Le « *clustering* » est par exemple, basé sur le principe du regroupement d'éléments ayant des valeurs de primitives les plus proches ou similaires. Certaines méthodes ajoutent également une contrainte pour éviter que des éléments ayant des valeurs de primitives éloignées ou dissimilaires, se retrouvent dans les mêmes groupes. La technique SVM dite *Support Vector Machine*, est basée sur des fonctions "noyau" qui ne sont que des formes de mesures de similarité. Par exemple, la fonction noyau gaussienne entre deux valeurs d'une primitive x_1 et x_2 est définie comme $K(x_1, x_2) = \exp(-\alpha(\|x_1 - x_2\|))$ avec $\alpha > 0$. Etant donné que la valeur de cette fonction diminue si x_1 est proche de x_2 , alors cette fonction peut être interprétée comme une mesure de similarité [123]. Notons aussi que la règle de décision par vote majoritaire (cas de la fusion de plusieurs décisions) ou la règle de classification basée sur la similarité entre l'objet à classer et un ensemble de cas d'apprentissage (par exemple, l'algorithme k-NN) sont fondées sur l'utilisation du concept de similarité.

Dans ce chapitre, nous allons présenter, dans un premier temps, le concept de similarité entre deux informations ainsi que les catégories existantes des mesures de similarité qui sont proposées dans la littérature. L'attention sera focalisée sur la similarité entre des informations possibilistes (i.e. la similarité entre distributions de possibilités) encodant des connaissances liées à la description pixelique de classes thématiques présentes dans une image. Face à la difficulté souvent rencontrée pour le choix d'une mesure de similarité, nous proposons ensuite une méthode de sélection d'une mesure de similarité qui soit la plus en adéquation afin de réaliser l'opération de classification pixelique. Cette méthode est basée sur le choix de la mesure permettant une meilleure discrimination entre les classes thématiques dans la scène.

Au-delà de l'application classique du concept de similarité pour la classification pixelique, nous proposons dans la dernière partie de ce chapitre, une démarche méthodologique permettant de réaliser l'opération de démixage pixelique en se basant sur le concept de similarité possibiliste spatiale. Cette démarche sera appliquée et évaluée dans le contexte d'un mono-capteur, là où le démixage n'est presque jamais appliqué. L'une des forces de la démarche proposée est d'identifier de manière efficace le contenu informationnel de régions ayant des modalités complexes ou bien composées de pixels mixtes (des pixels comportant un mélange de plus d'une classe parmi les différentes classes thématiques contenues dans la scène analysée).

4.2 Mesures de similarité

Tout système ayant pour but d'analyser ou d'organiser automatiquement un ensemble de données ou de connaissances (système de fouille des données, par exemple) doit utiliser sous une forme ou une autre, une mesure de similarité dont le but est d'établir les ressemblances ou les relations qui existent entre les informations manipulées [130]. Cette mesure fournit la possibilité d'ordonner les données, de les hiérarchiser ou de les classer. Une mesure de similarité « Sim » entre deux informations X et Y définies sur un univers de discours Ω (par exemple, deux objets représentés par deux vecteurs d'attributs, un objet et un ensemble d'objets ou deux ensembles d'objets) est une fonction qui prend ses valeurs dans l'intervalle $[0,1]$ et satisfait les propriétés suivantes, $\forall X, Y \in \Omega$:

- *Non négativité* : $\text{Sim}(X, Y) \geq 0$,
- *Symétrie* : $\text{Sim}(X, Y) = \text{Sim}(Y, X)$,
- *Autosimilarité maximale* : $\text{Sim}(X, X) = 1$.

En fonction de la nature des données à comparer, plusieurs mesures de similarité ont été proposées dans la littérature [124]. La conception de ces mesures est basée essentiellement sur deux points de vue : géométrique et ensembliste [121].

Les mesures basées sur le point de vue géométrique, appelées *mesures de similarité métriques*, considèrent les informations à comparer comme étant des points dans un espace métrique de représentation. Par conséquent, ces mesures réalisent dans un premier temps, une mesure de distance entre les informations à comparer et transforment ensuite, cette mesure de distance en une mesure de similarité.

En revanche, les mesures de similarité ensemblistes considèrent chaque information représentant un objet, comme étant un sous-ensemble algébrique de l'univers des primitives Ω (comportant les primitives réalisées par cette information). Par conséquent, les mesures de similarité ensemblistes sont basées sur le nombre de primitives communes ou distinctes, aux informations considérées (représentant les objets à comparer).

Sur la base de ces deux catégories de mesures de similarité, plusieurs travaux de recherche ont proposé des généralisations afin de réaliser la comparaison entre informations ambiguës représentées comme des ensembles flous définis sur l'univers Ω [121]. Les mesures de similarité ainsi obtenues sont appelées mesures de similarité floues. Dans la suite, nous présentons brièvement ces trois catégories de mesures : les mesures de similarité métriques, ensemblistes et floues.

Dans le cas où les informations à comparer sont de nature possibiliste (i.e. chaque information est représentée par une distribution de possibilités), l'immense majorité des travaux existants propose de transposer les mesures de similarité floues en substituant les fonctions d'appartenance aux ensembles flous par les distributions de possibilités. Malgré le fossé sémantique en terme de différence de signification physique entre une fonction d'appartenance (représentant un ensemble flou) et une distribution de possibilités (encodant la possibilité que chaque singleton soit l'unique qui s'est produit), nous n'allons pas mettre en cause cette transposition et nous considérons donc, ce même point de vue.

4.2.1 Mesures de similarité métriques

Ces mesures sont utilisées dans divers domaines pour quantifier la ressemblance entre deux informations (i.e. deux objets) représentées par deux vecteurs de primitives numériques appartenant à un espace métrique de dimension finie (souvent, $\Omega = \mathbb{R}^N$). Dans ce cas, la similarité entre deux informations est considérée comme étant inversement proportionnelle à une mesure de distance entre les deux vecteurs de primitives. Rappelons, à ce stade, qu'une fonction de mesure de distance $d(.,.)$ est définie par

$$\begin{aligned} d : \Omega \times \Omega &\rightarrow \mathbb{R}^+ \\ (X, Y) &\rightarrow d(X, Y) \end{aligned}$$

où $d(.,.)$ vérifie les axiomes d'une métrique :

- *Autosimilarité constante* : $d(X, X) = d(Y, Y), \forall X, Y \in \Omega$;
- *Minimalité* : $d(X, X) \leq d(X, Y), \forall X, Y \in \Omega$;
- *Symétrie* : $d(X, Y) = d(Y, X), \forall X, Y \in \Omega$;
- *Inégalité triangulaire* : $d(X, Y) \leq d(X, Z) + d(Z, Y), \forall X, Y \text{ et } Z \in \Omega$.

Sur la base d'une telle mesure de distance, une mesure de similarité est obtenue par la transformation de cette mesure via une fonction de proportionnalité inverse $\text{Sim}(X, Y) = g[d(X, Y)]$ pour laquelle $g(0) = 1$. Alors que la distance mesure le degré de différence entre deux vecteurs, une mesure de similarité donne

le degré de ressemblance entre ces deux vecteurs. Deux vecteurs ayant une grande similarité seront proches en termes de distance et vice versa. Par conséquent, la recherche des informations similaires se traduit par la recherche de faibles distances ou des mesures de similarité élevées.

Comme exemple de ces mesures, nous citons les distances de *Minkowski* [125] qui sont les plus souvent utilisées. Etant donnés deux vecteurs X et Y de dimension N , $X = [x_1, x_2 \dots x_N]^T$ et $Y = [y_1, y_2 \dots y_N]^T$, les équations (4.1) et (4.2) indiquent les distances de Minkowski et une mesure de similarité associée à cette distance :

$$d_p(X, Y) = \sqrt[p]{\sum_{i=1}^N |x_i - y_i|^p} \quad (4.1)$$

$$\text{Sim}_p(X, Y) = 1 - \frac{d_p(X, Y)}{\sqrt[p]{N}} \quad (4.2)$$

où $p \in [-\infty, +\infty]$ est appelé facteur d'ordre de la distance de Minkowski. Trois valeurs de p sont souvent utilisées :

La distance de Manhattan ($p = 1$) :

$$d_M(X, Y) = \sum_{n=1}^N |x_n - y_n|$$

La distance Euclidienne ($p = 2$) :

$$d_E(X, Y) = \sqrt{\sum_{n=1}^N |x_n - y_n|^2}$$

La distance de Chebychev ($p = \infty$) :

$$d_{Ch}(X, Y) = \max |x_n - y_n|, n = 1, 2, \dots, N.$$

4.2.2 Mesures de similarité ensemblistes

Les mesures de similarité ensemblistes considèrent l'univers Ω des singletons comme étant les primitives pouvant se produire par une information. Avec cette vision, une information décrivant un objet peut être représentée par un sous-ensemble de primitives de Ω . Par conséquent, mesurer la similarité (i.e. quantifier la ressemblance) entre deux informations revient à comptabiliser le « nombre » de primitives communes (ou distinctes) entre les deux objets représentés par ces informations.

Le modèle proposé par Tversky [131] est la mesure de similarité ensembliste la plus utilisée. En supposant A et B deux sous-ensembles de primitives représentant deux objets, la mesure de similarité de Tversky est définie de la façon suivante :

$$S_{\alpha, \beta}(A, B) = \frac{|A \cap B|}{|A \cap B| + \alpha |A - B| + \beta |B - A|} \quad (4.3)$$

avec :

$|A \cap B|$ le cardinal du sous ensemble d'intersection entre A et B (i.e. nombre de primitives communes contenues dans les deux informations) ;

$|A-B|$ (resp. $|B-A|$) désigne le cardinal du complémentaire de B dans A (resp. de B dans A), et

$\alpha, \beta \geq 0$ sont deux facteurs de pondération permettant de différencier l'importance des primitives dans les deux informations considérées.

Plusieurs mesures particulières de similarité ensemblistes (correspondant à des cas particuliers du modèle de Tversky) sont proposées dans la littérature.

Par exemple, l'*indice de Jaccard* est la mesure ensembliste correspondant aux valeurs particulières de $\alpha = \beta = 1$:

$$S_{\alpha=1, \beta=1}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Cette mesure de similarité ensembliste correspond à la vision la plus simple qui consiste à considérer la similarité comme étant le rapport entre le nombre de primitives communes et le nombre total de primitives.

Pour $\alpha = 1$ et $\beta = 0$ (resp. $\alpha = 0$ et $\beta = 1$) la mesure de similarité ensembliste de Tverski se réduit à la *mesure d'inclusion* de B dans A (resp. A dans B) :

$$\begin{aligned} \text{Sim}_{\text{Tversky}}(A, B) |_{\alpha=1, \beta=0} &= \frac{f(A \cap B)}{f(A)} \\ \text{Sim}_{\text{Tversky}}(A, B) |_{\alpha=0, \beta=1} &= \frac{f(A \cap B)}{f(B)} \end{aligned}$$

Finalement, pour $\alpha = \beta = 0.5$, la mesure de similarité ensembliste de Tverski porte le nom du *coefficient de Dice* :

$$\text{Sim}_{\text{Tversky}}(A, B) |_{\alpha=\beta=0.5} = \frac{2f(A \cap B)}{f(A) + f(B)}$$

4.2.3 Mesures de similarité floues

Les mesures de similarité floues constituent une généralisation des mesures de similarité métriques et ensemblistes lorsque les informations décrivant les objets sont de nature floue (i.e. l'information décrivant un objet par l'intermédiaire d'un ensemble flou défini sur l'univers des primitives Ω).

De nombreuses contributions de mesures de similarité floues ont déjà été proposées [132] et appliquées dans des domaines tels que le traitement d'image, la reconnaissance de formes ou la fouille de données [61][133][134].

A titre d'exemple, en supposant $\Omega = \{x_1, x_2, \dots, x_N\}$ l'univers des primitives, la distance de *Minkowski* d'ordre p peut être généralisée aux informations floues A et B définie via les fonctions d'appartenance $\mu_A(\cdot)$ et $\mu_B(\cdot)$ par :

$$d_{M,p}(A, B) = \left[\sum_{n=1}^N |\mu_A(x_n) - \mu_B(x_n)|^p \right]^{1/p}, \quad p \geq 1$$

Alors, en considérant la fonction de transformation distance-similarité

$$\text{Sim}(A, B) = g[d(A, B)] = 1 - d(A, B)$$

nous obtenons les mesures de similarité floues suivantes :

Mesure de Manhattan, ou W-Similarité, $p = 1$:

$$\text{Sim}_M(A, B) = 1 - \frac{1}{N} \sum_{n=1}^N |\mu_A(x_n) - \mu_B(x_n)|$$

Mesure de similarité euclidienne floue, $p = 2$,

$$\text{Sim}_E(A, B) = 1 - \sqrt{\frac{1}{N} \sum_{n=1}^N |\mu_A(x_n) - \mu_B(x_n)|^2}$$

Mesure de similarité de Chebychev, ou L-Similarité, $p = \infty$,

$$\text{Sim}_{Ch}(A, B) = 1 - \max_{n=1, \dots, N} |\mu_A(x_n) - \mu_B(x_n)|$$

Bouchon-Meunier *et al* [135] ont proposé une mesure de similarité floue qui est une extension du modèle de Tversky [131] et ont mis l'accent dans [136], sur l'importance du classement « *ranking* » induit par l'application d'une mesure de similarité donnée sur des paires d'objets plutôt que sur les valeurs de similarités elles-mêmes.

A titre d'exemple, l'indice de Jaccard flou est obtenu en substituant les ensembles algébriques par les ensembles flous :

$$\text{Sim}_M(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{\sum_{n=1}^N \min(\mu_A(x_n), \mu_B(x_n))}{\sum_{n=1}^N \max(\mu_A(x_n), \mu_B(x_n))}$$

Dans [151], la similarité entre deux ensembles flous a été interprétée comme une relation floue définie sur ces ensembles et qui correspond au degré avec lequel les deux ensembles flous sont égaux.

Notons que malgré les nombreuses mesures de similarité floues proposées dans la littérature, ces mesures ont été utilisées de façon arbitraire et empirique, sans justification ou argumentation convaincante [127]. En réalité, pour une application donnée, il est difficile de donner une justification satisfaisante pour l'utilisation de telle ou telle mesure de similarité choisie parmi une liste très longue de mesures. Une mesure de similarité peut être très efficace pour une application donnée et entièrement infructueuse dans un autre contexte [128]. Cela montre la nécessité de proposer une méthodologie permettant d'avoir un consensus pour le choix d'une telle ou telle mesure de similarité dans une application donnée.

Comme nous l'avons déjà mentionné auparavant, dans le cas où les informations décrivant des objets sont de nature possibiliste, l'immense majorité des travaux de recherche généralise les mesures de similarité floues aux informations possibilistes en remplaçant les fonctions d'appartenance aux ensembles flous par les distributions de possibilités. Dans la suite de ce chapitre, nous allons présenter l'intérêt du concept de similarité dans le processus d'interprétation de scène tout en proposant une méthodologie de sélection d'une mesure de similarité pour la discrimination entre les classes thématiques dans la scène.

4.3 Similarité et interprétation de scène

Comme nous avons souligné précédemment, dans tout processus d'interprétation de scène, les informations disponibles peuvent être très complexes, hétérogènes (quantitatives, qualitatives, binaires, ordinales, ...) [126] et touchées par différents types d'imperfections (incertitude, imprécision, ambiguïté, ...) [121]. D'une manière générale, dans un processus d'analyse de scène, l'attribution d'un pixel, région ou objet, à une classe est basée sur une mesure de similarité entre l'information décrivant ce pixel (région ou objet) et les informations décrivant les différentes classes thématiques. De même, étudier la difficulté de discrimination revient à mesurer la similarité entre ces différentes classes. On peut donc

facilement constater l'importance majeure des mesures de similarité dans presque toutes les tâches réalisées au sein des systèmes d'interprétation de scène.

Par conséquent, la nature de l'imperfection qui affecte les différentes informations ainsi que la modélisation mathématique pour la représentation de ces informations est d'une importance cruciale car ceci conditionne les différentes opérations à réaliser sur ces informations et, en particulier, la nature des mesures de similarité à utiliser.

Nous rappelons que l'ambiguïté est due principalement à la qualification par des termes linguistiques des caractéristiques du système d'interprétation, ou à la faible représentativité des bases d'exemples. A titre d'exemple, pour caractériser la classe « micro-calcification » dans les images mammographiques, un radiologue utilise des termes linguistiques tels que brillant, rond, lisse, long, etc. [154]. Une source supplémentaire de l'ambiguïté, qui est spécifique à la classification pixelique, concerne l'existence de pixels mixtes (i.e. des pixels représentant des cellules de résolution comportant plusieurs classes thématiques en même temps). Ce type d'informations constitue un problème récurrent dans toute tentative de discrimination interclasses dans un processus d'interprétation de scène. Généralement, plus les classes sont « faciles » à discriminer (i.e. moins de confusion interclasses), plus elles peuvent être représentées par des termes linguistiques qui sont sémantiquement différents [137].

Pour toutes ces raisons, la modélisation possibiliste des informations utilisées en interprétation de scène est tout à fait légitime. Ainsi, dans le cadre de la théorie des possibilités, la comparaison des informations imparfaites revient à mesurer la similarité entre des distributions de possibilités et évaluer ainsi l'ambiguïté due au « chevauchement » entre les classes. Plusieurs mesures de similarité ont été proposées à cet égard dans la littérature [127]. Néanmoins, il existe un réel besoin pour un consensus sur le choix d'une mesure de similarité [136].

Dans le paragraphe suivant, nous proposons une méthodologie de sélection d'une mesure de similarité dans le cadre particulier de la classification pixelique.

4.4 Similarité inter-classes

Considérons une scène comportant M classes thématiques $\{C_1, C_2, \dots, C_M\}$ et supposons $\Omega = \{x_1, x_2, \dots, x_N\}$ l'univers de discours représentant l'ensemble des valeurs potentielles de la primitive mesurée. Nous supposons aussi que chaque classe thématique C_m , $m=1, 2, \dots, M$ est décrite par une information possibiliste encodée par une distribution de possibilités π_m , $m=1, \dots, M$ définie sur Ω .

Dans ce contexte, et en considérant une mesure de similarité possibiliste $\text{Sim}(\dots)$, nous proposons de caractériser cette mesure par une *matrice de similarité* de dimensions $M \times M$ donnée par :

$$\text{Sim} = \begin{pmatrix} \text{Sim}(\pi_1, \pi_1) & \text{Sim}(\pi_1, \pi_2) & \dots & \text{Sim}(\pi_1, \pi_M) \\ \text{Sim}(\pi_2, \pi_1) & \text{Sim}(\pi_2, \pi_2) & \dots & \text{Sim}(\pi_2, \pi_M) \\ \dots & \dots & \dots & \dots \\ \text{Sim}(\pi_M, \pi_1) & \text{Sim}(\pi_M, \pi_2) & \dots & \text{Sim}(\pi_M, \pi_M) \end{pmatrix} \quad (4.4)$$

Cette matrice permet de nous renseigner, d'une façon synthétique, sur la similarité entre les différentes classes thématiques présentes dans une scène analysée. Il s'agit d'une matrice symétrique (car : $\text{Sim}(\pi_{m1}, \pi_{m2}) = \text{Sim}(\pi_{m2}, \pi_{m1})$, $\forall m1, m2 = 1, 2, \dots, M$) pour laquelle les éléments se trouvant sur la diagonale sont unitaires (car : $\text{Sim}(\pi_m, \pi_m) = 1$, $\forall m = 1, 2, \dots, M$).

Parmi les mesures de similarité les plus fréquemment rencontrées dans la littérature, les cinq mesures suivantes ont été retenues dans le cadre de notre étude :

- Trois mesures de *Minkowski* : étant donné que les distributions de possibilités sont représentées par des vecteurs des valeurs réelles dans l'intervalle $[0,1]$, les mesures de *Minkowski* peuvent être utilisées [129]. Les trois cas particuliers de mesures de Minkowski sont retenus :

1. Mesure de Manhattan :

$$\text{Sim}_1(\pi_1, \pi_2) = 1 - \frac{1}{N} \sum_{i=1}^N |\pi_1(x_i) - \pi_2(x_i)| \quad (4.5)$$

2. Mesure de similarité euclidienne floue :

$$\text{Sim}_2(\pi_1, \pi_2) = 1 - \sqrt{\frac{1}{N} \sum_{i=1}^N |\pi_1(x_i) - \pi_2(x_i)|^2} \quad (4.6)$$

3. Mesure de similarité de Chebychev :

$$\text{Sim}_\infty(\pi_1, \pi_2) = 1 - \max_{i=1}^N |\pi_1(x_i) - \pi_2(x_i)| \quad (4.7)$$

- La mesure de similarité IA « *Information Affinity* » : cette mesure de similarité entre deux distributions π_1 et π_2 , est basée sur la combinaison entre une mesure de distance $d(\pi_1, \pi_2)$ et une valeur d'inconsistance de l'intersection entre ces deux distributions $\text{Inc}(\pi_1, \pi_2)$ [141].

$$\text{Sim}_{\text{IA}}(\pi_1, \pi_2) = 1 - \frac{\kappa \cdot d(\pi_1, \pi_2) + \lambda \cdot \text{Inc}(\pi_1, \pi_2)}{\kappa + \lambda} \quad (4.8)$$

avec $\kappa > 0$ and $\lambda > 0$, $d(\pi_1, \pi_2)$ représente une distance entre les deux distributions de possibilités (la distance euclidienne de *Minkowski* entre les deux distributions π_1 et π_2 a été considérée).

Notons que cette mesure de similarité possibiliste est parmi les très rares mesures de similarité possibiliste qui ne soit pas une généralisation directe des mesures de similarité floues au cas possibiliste.

- La mesure de similarité SI « *Similarity Index* » : cette mesure entre les distributions π_1 et π_2 , est définie par [141]:

$$\text{Sim}_{\text{SI}}(\pi_1, \pi_2) = \min\{\alpha(\pi_1, \pi_2), \alpha(1-\pi_1, 1-\pi_2)\} \quad (4.9)$$

où

$$\alpha(\pi_1, \pi_2) = \frac{\sum_{i=1}^N \pi_1(x_i) \times \pi_2(x_i)}{\sum_{i=1}^N \{\max(\pi_1(x_i), \pi_2(x_i))\}^2} \quad (4.10)$$

si $\sum \{\max(\pi_1(x_i), \pi_2(x_i))\}^2 = 0$, alors $\text{Sim}_{\text{SI}}(\pi_1, \pi_2) = \alpha(\pi_1, \pi_2) = 1$.

La question pour laquelle nous allons essayer d'apporter une réponse est : comment sélectionner la mesure de similarité la mieux adaptée pour la discrimination des classes dans un contexte d'interprétation de scène ?

Dans cet objectif, nous proposons de considérer comme critère de sélection, la distance euclidienne D_E entre les matrices de similarité (correspondant aux différentes mesures de similarité possibilistes) et la matrice unité I_M (qui est supposée correspondre à la mesure de similarité optimale) :

$$D_E = \sqrt{\sum_{i,j \in \{1,2,\dots,M\}} [\text{Sim}(\pi_i, \pi_j) - I_M(i,j)]^2} \quad (4.11)$$

Plus la distance D_E est faible, meilleure est la discrimination inter-classe de la mesure de similarité évaluée. Par conséquent, la mesure de similarité possibiliste à retenir est celle dont la matrice de similarité associée est la plus proche de la matrice unité I_M au sens de la distance euclidienne.

4.4.1 Exemple d'évaluation de la similarité entre deux classes

Pour évaluer l'approche de sélection de la mesure de similarité, nous proposons de considérer le cas d'une scène comportant deux classes C_1 et C_2 . Les pixels de ces deux classes C_1 et C_2 suivent dans un premier temps, un loi de probabilité gaussienne $G(m, \sigma)$ (avec une moyenne « m » et un écart type « σ ») sur laquelle la transformation symétrique de Dubois-Prade a été appliquée et dans un deuxième temps, un loi de probabilité Rayleigh $R(\alpha)$ avec un paramètre α sur laquelle la transformation symétrique de Dubois-Prade a été aussi appliquée.

La raison de ce choix réside dans le fait que ces deux types de lois sont les plus utilisés en imagerie pour la modélisation des classes existantes dans la scène. Les valeurs initiales choisies pour les paramètres caractéristiques de chaque classe sont les suivantes : pour la classe C_1 , $m_1=120$, $\sigma_1=20$; pour la classe C_2 , $m_2=140$, $\sigma_2=20$ (Figure 4.1-(a) et (c)) (dans le cas des distributions gaussiennes) et $\alpha_1 = \alpha_2 = 40$ (dans le cas des distributions Rayleigh (Figure 4.2-(a))).

Pour étudier le comportement des mesures de similarité proposées, la distance euclidienne D_E a été calculée pour chaque mesure de similarité en faisant varier, dans un premier temps, la moyenne de la classe C_2 puis, dans un deuxième temps, l'écart-type de la classe C_2 générée selon la densité de probabilités gaussienne (resp. le paramètre α_2 de la classe C_2 générée selon la loi de Rayleigh), tout en conservant une valeur fixe pour la moyenne et l'écart-type de la classe C_1 (resp. une valeur fixe pour le paramètre α_1 de la classe C_1).

Les Figure 4.1-(b) et -(d) présentent l'évolution de la mesure de la distance euclidienne D_E au fur et à mesure que les moyennes s'éloignent ou que les écarts-types diffèrent.

La Figure 4.2-(b) présente l'évolution de la mesure de la distance euclidienne D_E au fur et à mesure que les paramètres α_1 , α_2 s'éloignent.

D'après les courbes de la Figure 4.1-(b) et -(d) et la Figure 4.2-(b), la mesure de la similarité de Chebychev $\text{Sim}_\infty(\pi_1, \pi_2)$ est la mesure de similarité possibiliste la plus adaptée parmi les cinq mesures choisies pour décrire la similarité entre les deux classes gaussiennes. En effet, cette mesure est celle qui produit la distance euclidienne la plus faible, parmi les différentes mesures testées, entre la matrice de similarité et la matrice unité et ceci lorsque les valeurs étudiées de $|m_2 - m_1|$, $|\sigma_2 - \sigma_1|$, et $|\alpha_2 - \alpha_1|$ augmentent.

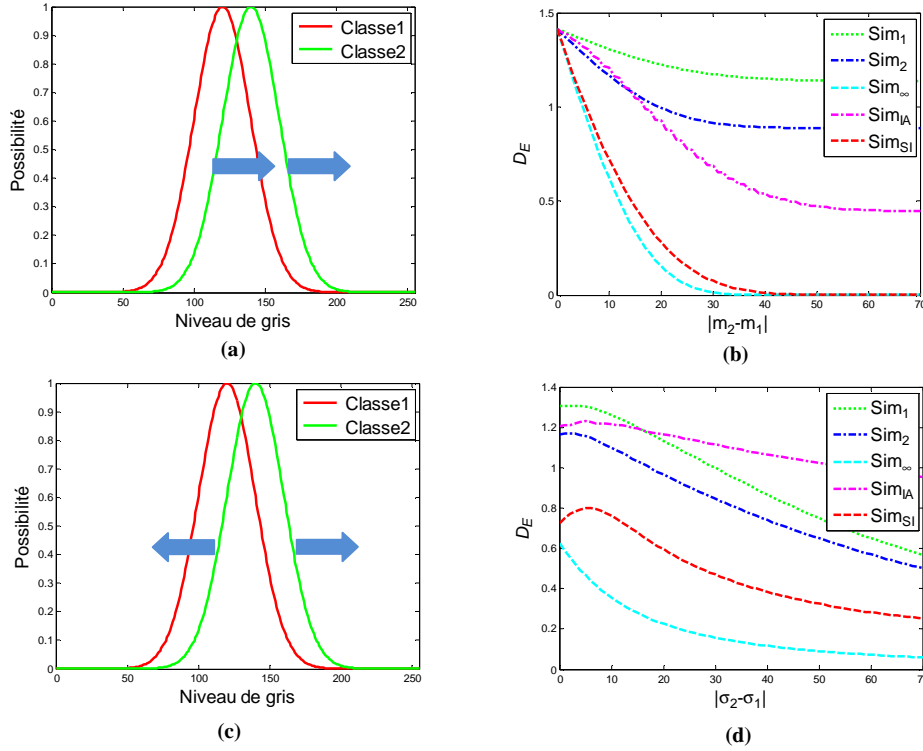


Figure 4.1 : Distance euclidienne D_E correspondant aux différentes mesures de similarité possibilistes, (a) et (c) deux distributions de possibilités gaussiennes représentant les deux classes C_1 et C_2 , (b) évolution de la mesure D_E en fonction de la différence des moyennes entre les classes C_1 et C_2 , (d) évolution de la mesure D_E en fonction de la différence d'écart-types entre les classes C_1 et C_2 .

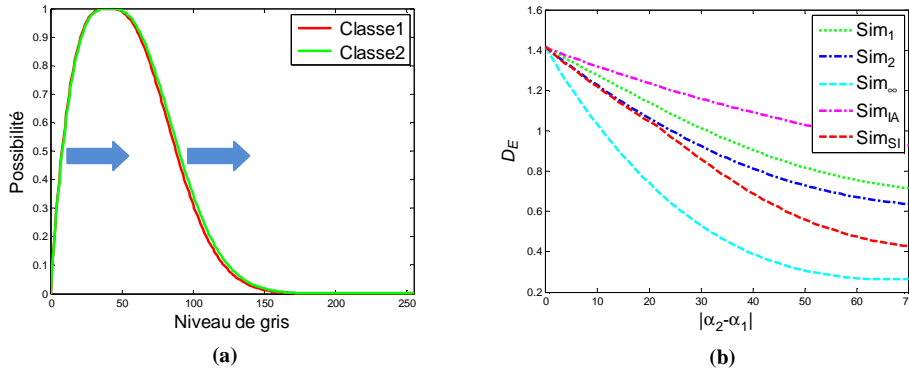


Figure 4.2 : Distance euclidienne D_E correspondant aux différentes mesures de similarité possibilistes, (a) deux distributions de possibilités de Rayleigh représentant les deux classes C_1 et C_2 , (b) évolution de la mesure D_E en fonction de la différence de paramètre α_1 et α_2 entre les classes C_1 et C_2 .

Il est important de noter que la mesure de similarité possibiliste de Chebychev, Sim_∞ , est la seule mesure de similarité (parmi les 5 mesures considérées) qui ne fait pas intervenir toutes les valeurs des primitives de l'univers Ω . En effet, cette mesure ne fait intervenir que la valeur de primitive pour laquelle l'écart des degrés de possibilités est le plus élevé entre les distributions de possibilités considérées.

Dans la suite nous allons retenir cette mesure de similarité Sim_∞ dans l'approche proposée pour la réduction d'ambiguïté entre les classes et ceci dans un cadre de représentation possibiliste. Deux exemples d'applications liées à l'interprétation de scènes, des mesures de similarité possibilistes seront abordés dans les paragraphes suivants : la classification et le démixage pixelique.

4.5 Similarité possibiliste et classification pixelique

Après avoir présenté le concept de similarité et les différentes mesures de similarité ainsi qu'une démarche permettant de sélectionner la mesure de similarité possibiliste la mieux adaptée pour la discrimination de classes thématiques en vue d'une classification pixelique, nous allons proposer dans ce paragraphe une démarche de classification pixelique basée sur les mesures de similarité possibilistes. En effet, nous supposons d'opérer la classification pixelique sous l'hypothèse du monde fermé. Cette hypothèse signifie que l'image analysée I représente une scène analysée composée d'un ensemble exclusif et exhaustif $\psi = \{C_1, C_2, \dots, C_M\}$ comportant M classes thématiques prédéfinies.

4.5.1 Démarche de classification pixelique

Le processus « classique » de classification pixelique possibiliste consiste à représenter les connaissances *a priori* sur les M classes par l'intermédiaire de M distributions de possibilités $\pi_{C_m}(\cdot)$, $m = 1, \dots, M$, définies sur l'univers des primitives Ω . Dans un deuxième temps, et en adoptant la règle de décision de la possibilité maximale, chaque pixel $P \in I$ de l'image analysée pour lequel la valeur de primitive observée est $x(P)$, sera attribué à la classe ayant la valeur de possibilité maximale :

$$\text{Classe}(P) = C_{m_0} \quad \text{Si et Seulement Si} \quad \pi_{C_{m_0}}(x(P)) = \max_{m=1}^M \pi_{C_m}(x(P))$$

Notons que ce processus de classification peut être qualifié de ponctuel car la prise de décision de classification ne fait intervenir que la valeur de la primitive mesurée au niveau de chaque pixel sans exploiter le contexte spatial (qui d'ailleurs peut réduire l'ambiguïté potentielle entre les différentes classes). Une deuxième difficulté de ce processus de classification réside dans le fait que les distributions de possibilités utilisées représentent des connaissances plus faibles que les densités de probabilités car ces distributions sont issues soit d'une description par l'expert, des différentes classes thématiques via des variables linguistiques des primitives observées, soit d'une caractérisation non représentatives par l'expert de ces classes en indiquant des zones réduites de pixels "purs" appartenant à ces différentes classes. Dans ce contexte, et afin de remédier à ces deux lacunes rencontrées en classification pixelique possibiliste, nous proposons d'utiliser l'hypothèse intuitive que l'immense majorité des pixels se trouve dans des régions homogènes. Comme illustré sur la Figure 4.3, sur la base de cette hypothèse, la démarche de classification possibiliste proposée pour chaque pixel $P \in I$ consiste à :

- Considérer un voisinage spatial $V(P)$, par exemple un voisinage de 3×3 pixels ou 5×5 pixels, centré autour de chaque pixel P ;
- Estimer la distribution de possibilités $\pi_P(\cdot)$ de la classe à laquelle sont supposés appartenir tous les pixels de ce voisinage $V(P)$;
- Réaliser une classification pixelique en se basant sur les valeurs mesurant la similarité possibiliste $\text{Sim}_\infty(\pi_P, \pi_{C_m})$, $m = 1, \dots, M$ pour la prise de décision de classification possibiliste (et ceci à la place des valeurs de possibilités ponctuelles $\pi_{C_m}(P)$, $m = 1, \dots, M$).

La méthode d'estimation de la distribution de possibilités est celle proposée dans le chapitre précédent à partir d'une information limitée fournie par l'expert au travers des zones d'apprentissage

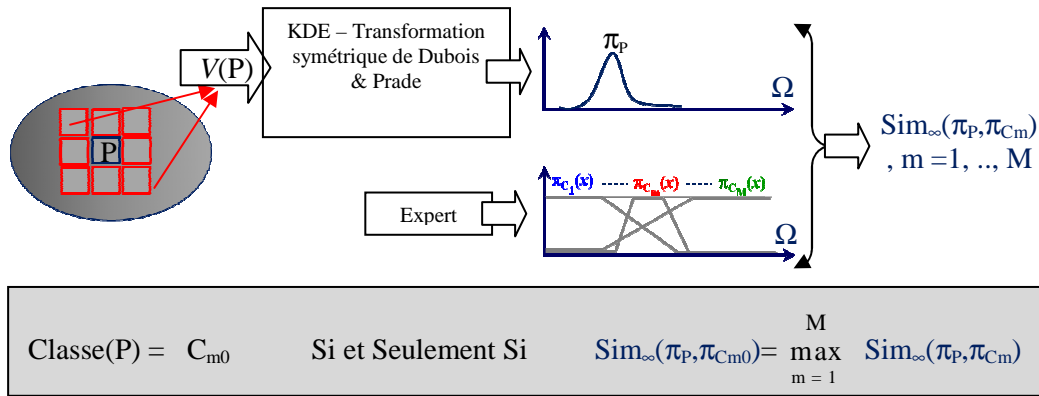
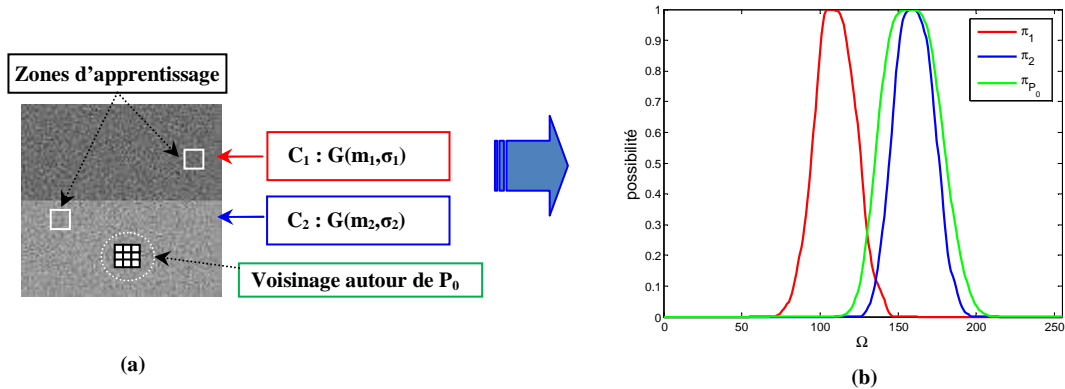


Figure 4.3 : Similarité possibiliste et classification pixelique.

La mesure de similarité utilisée dans cette étape est la mesure de similarité Sim_{∞} retenue dans l'étape d'évaluation de la similarité entre les classes. La Figure 4.4 illustre l'approche proposée dans le cas d'une image de synthèse de taille 100×100 pixels et composée de deux classes. L'intensité des pixels dans chaque classe a été générée par une distribution normale. Les valeurs choisies pour les paramètres caractéristiques de chaque classe sont les suivantes : pour la classe C_1 , $m_1=110$, $\sigma_1=10$; pour la classe C_2 , $m_2=160$, $\sigma_2=10$.

Figure 4.4 : (a) Image de synthèse originale, zones d'apprentissage et une zone autour du pixel considéré P_0 ; (b) distributions de possibilités π_1 , π_2 , et π_{P_0} .

La Figure 4.4 montre que la caractérisation de chaque pixel P_0 de l'image I est réalisée en prenant en compte son voisinage au travers de la constitution de distribution de possibilités $\pi_{P_0}(\cdot)$. En effet, l'approche proposée permet de comparer cette distribution de possibilités locale avec les distributions de possibilités des différentes classes, et non pas seulement un degré d'appartenance de paramètres extraits du voisinage du P_0 aux distributions de possibilités des classes.

4.5.2 Résultats de classification pixelique

En utilisant des images de synthèse ainsi que des images réelles, nous allons évaluer les performances de l'approche proposée. Nous proposons de comparer les résultats obtenus avec ceux des approches de classification basées sur les champs de Markov [63] [68]. Rappelons que les champs de Markov constituent un outil utile et performant pour caractériser l'information relative au contexte spatial. De plus, les approches de classification utilisant les champs de Markov sont considérées comme étant la

référence pour ce type de problèmes de classification. La méthode de type markovienne que nous choisissons, utilise l'algorithme ICM « *Iterated Conditional Modes* » [21] pour le calcul de la configuration optimale du champ d'étiquettes. En effet, le choix d'une telle approche basée sur l'algorithme ICM est dû au fait que cette approche est la plus fréquemment utilisée dans la classification contextuelle [18], et a prouvé son efficacité dans l'amélioration des résultats de classification.

4.5.2.1 Images simulées

Pour ce faire, l'image de synthèse de la Figure 4.4 est d'abord considérée pour l'évaluation des performances de l'approche proposée d'une part, et des performances de l'approche de classification basée sur les champs de Markov, d'autre part. Les distributions de possibilités sont estimées à partir des zones d'apprentissage délimitées par l'expert sur l'image I . De même, les distributions de possibilités locales $\pi_{p_0}(x)$ sont estimées à partir de fenêtres de taille 3×3 , 5×5 et 7×7 autour de chaque pixel de l'image de synthèse. En mesurant la similarité entre la distribution de possibilités locale π_{p_0} et chacune des M distributions de possibilités caractérisant les deux classes thématiques, chaque pixel est attribué à la classe pour laquelle la similarité est la plus élevée. Dans l'approche Markovienne, l'algorithme K-moyennes est utilisé pour la segmentation initiale de l'image de synthèse et l'algorithme ICM est itéré 15 fois, avec une 8-connexité et un paramètre d'isotropie β égal à 1.5. En effet, ces réglages de l'algorithme ICM ont permis d'avoir les meilleurs résultats de classification des images de télédétection [63] et des images obtenues par imagerie IRM [68]. L'optimalité de ces paramètres a été validée empiriquement sur l'image de synthèse.

Pour chaque expérience, le taux d'erreur de classification est calculé en faisant varier la différence des moyennes et des écarts-types entre les deux classes synthétisées. La Figure 4.5 présente les taux d'erreur dans le cas Markovien et dans le cas de l'approche proposée observés pour des fenêtres de taille 3×3 .

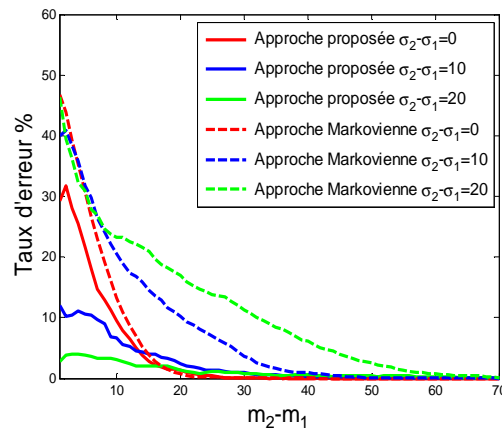


Figure 4.5 : Evolution du taux d'erreur de classification en fonction des caractéristiques des deux classes (moyennes et écarts-types) dans le cas Markovien (traits en pointillés) et dans le cas de l'approche proposée (traits pleins) pour des fenêtres de taille 3×3

Nous remarquons la supériorité des performances de l'approche proposée par rapport à l'approche markovienne. En effet, le taux d'erreur obtenu de la classification pixelique lorsqu'on utilise la méthode proposée est toujours moins élevé que le taux d'erreur obtenu par l'approche Markovienne. Notons aussi que l'approche proposée permet de réduire significativement le taux d'erreur quand le chevauchement entre les deux classes augmente (faibles différences d'écart-type pour des valeurs de moyennes proches).

La Figure 4.6 présente les taux d'erreur dans le cas de l'approche proposée observés pour des fenêtres de taille 3×3 , 5×5 et 7×7 .

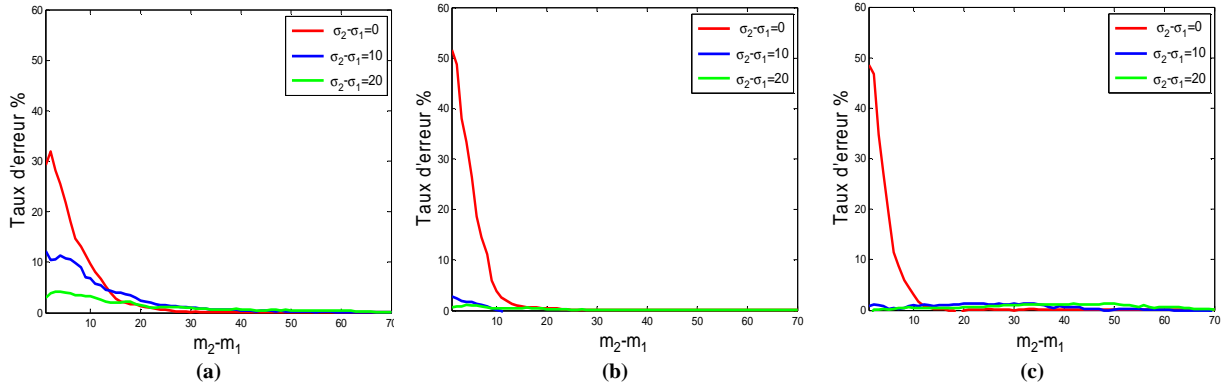


Figure 4.6 : Evolution du taux d'erreur de classification en fonction des caractéristiques des deux classes (moyennes et écarts-types) dans le cas de l'approche proposée pour des fenêtres de taille 3×3 (a), 5×5 (b) et 7×7 (c)

Nous constatons aussi que l'estimation des distributions de possibilités locales $\pi_{p0}(x)$ à partir des fenêtres de taille 5×5 et 7×7 ont permis d'avoir une meilleure réduction d'ambiguïté entre les classes par rapport aux fenêtres de taille 3×3. De plus, nous constatons que l'augmentation du voisinage de plus de 5×5 n'apporte pas une grande réduction de taux d'erreurs.

La Figure 4.7 montre les résultats en termes de taux d'erreur lorsque les deux classes thématiques sont engendrées par des densités de probabilité du type Rayleigh. Nous remarquons que ces résultats aboutissent aux mêmes conclusions que dans le cas des classes thématiques engendrées par des lois gaussiennes.

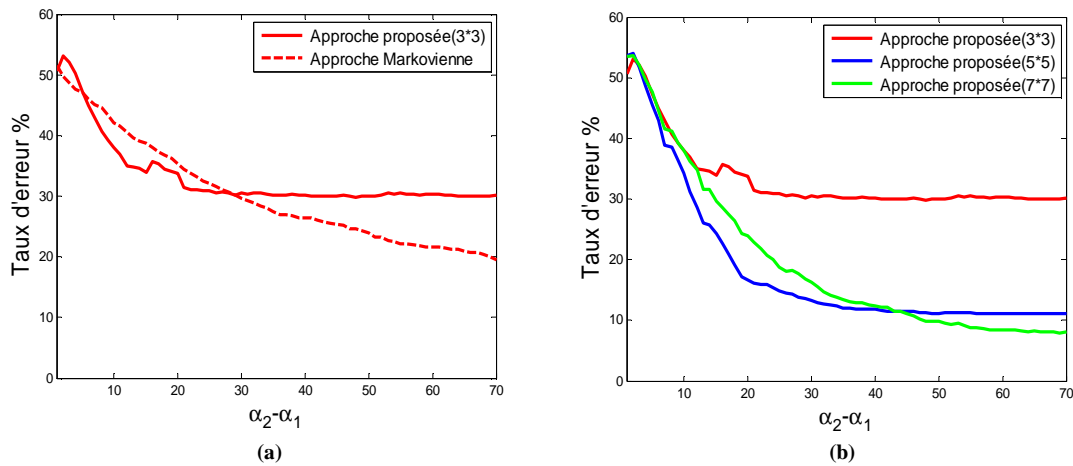


Figure 4.7 : Evolution du taux d'erreur de classification en fonction des caractéristiques des deux classes (α_1 et α_2) (a) dans le cas Markovien et dans le cas de l'approche proposée pour des fenêtres de taille 3×3 (b) dans le cas de l'approche proposée pour des fenêtres de taille 3×3, 5×5 et 7×7

Finalement, la Figure 4.8 présente un exemple de résultat de classification de l'image de synthèse par application de l'approche Markovienne ainsi que l'approche proposée (pour des fenêtres de taille 3×3) et pour les valeurs suivantes : $C_1(m_1 = 110, \sigma_1 = 10)$ et $C_2(m_2 = 120, \sigma_2 = 20)$ dans le cas gaussien et $\alpha_1 = 40$ et $\alpha_2 = 70$ dans le cas de Rayleigh.

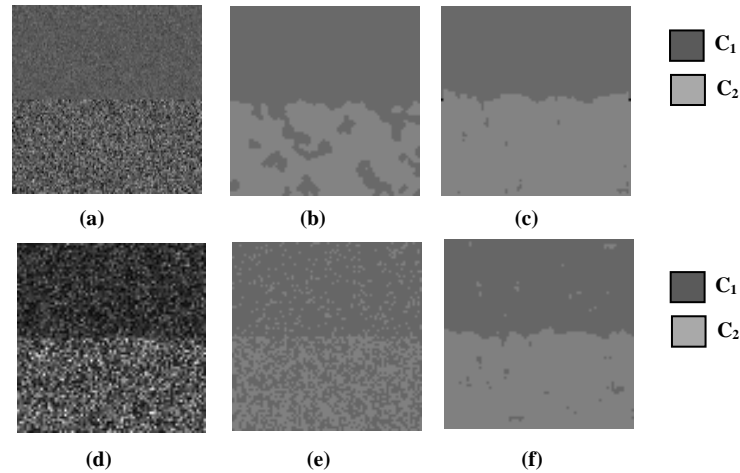


Figure 4.8 : Classification d'image de synthèse
 (a) image originale, cas gaussienne et (d) image originale, cas Rayleigh ;
 (b) et (e) résultats de la classification par l'approche Markovienne (cas gaussien et cas Rayleigh respectivement) et ;
 (c) et (f) résultats de la classification par l'approche proposée (cas gaussienne et cas Rayleigh respectivement)

On observe une homogénéité intéressante des régions ainsi déterminées à partir d'échantillons reposant sur des mesures limitées à des fenêtres de taille 3×3 et cette homogénéité est plus importante que celle obtenue par l'approche Markovienne.

Pour interpréter la supériorité des résultats obtenus par l'approche proposée par rapport à l'approche markovienne, il faut noter que les résultats obtenus par l'approche markovienne dépendent fortement de la classification initiale réalisée (que nous avons obtenu par l'application de l'algorithme des K-moyennes). Ce choix a été réalisé afin de comparer les deux approches dans un contexte de complexité modérée. Notre objectif n'est pas de montrer que l'approche proposée est meilleure que l'approche markovienne, mais qu'au moins les performances que nous obtenons sont comparables à l'approche de référence qui est celle des méthodes markoviennes. Notons que cette comparaison entre l'approche proposée et l'approche Markovienne se base sur des images de synthèse pour lesquelles la génération est réalisée sans tenir compte de la corrélation pixelique spatiale.

4.5.2.2 Images réelles

Nous allons évaluer les performances de l'approche de classification proposée sur des images réelles. En effet, la Figure 4.9 présente un exemple de classification d'une image « optique » d'une tumeur du type mélanome (image en niveau de gris et de taille 273×433 pixels et sans réalité terrain) par l'approche markovienne ainsi que par l'approche proposée. Pour l'approche markovienne, les deux distributions de probabilités conditionnelles $\Pr(x_n/C)$ associées aux deux classes C_1 (tumeur) et C_2 (peau saine) sont supposées de type gaussien et les autres paramètres sont les mêmes que ceux utilisés dans le cas de l'image de synthèse.

Pour l'approche proposée, la phase d'apprentissage a permis de caractériser ces deux classes C_1 et C_2 par des distributions de possibilités estimées à partir des zones d'apprentissage positionnées par l'expert. En utilisant les mesures de similarité Sim_∞ entre la modélisation possibiliste de chacune des deux classes et les distributions de possibilité locales associées à chaque pixel, les résultats de la classification sont présentés dans la Figure 4.9.

Ne disposant pas d'une réalité terrain, il est bien évident que nous ne pouvons pas réaliser la comparaison entre les approches via des paramètres quantitatifs du type taux d'erreur. Nous nous limitons, donc, à une simple comparaison visuelle. En effet, l'analyse visuelle des résultats de

classification nous permet d'observer une homogénéité intéressante des régions déterminées à partir des fenêtres de taille 3×3. Cette homogénéité est compatible avec celle obtenue par l'approche markovienne.

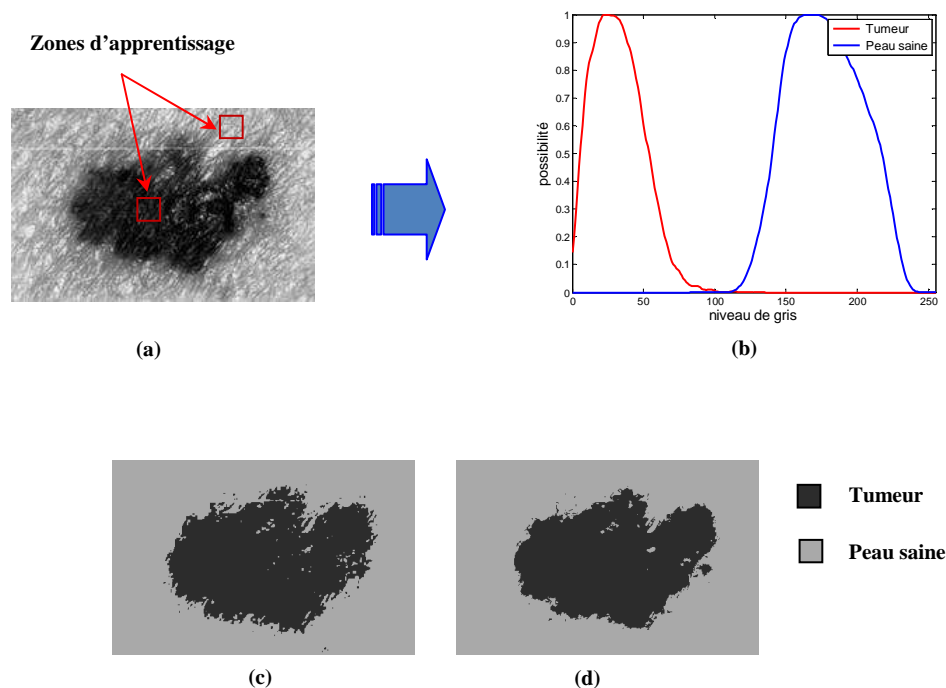


Figure 4.9 : Classification d'image de mélanome
 (a) image originale et zones d'apprentissage,
 (b) estimation des distributions de possibilités des deux classes considérées,
 (c) classification par la méthode markovienne;
 (d) classification par l'approche proposée

4.6 Similarité possibiliste et démixage pixelique

Un deuxième exemple d'application de la similarité possibiliste est présenté dans ce paragraphe. Il s'agit de la problématique de démixage pixelique. Dans un premier temps, nous allons expliciter cette application, puis nous proposons d'étudier l'application de la démarche de classification pixelique que nous venons de présenter, dans l'objectif de réaliser le démixage pixelique. Les résultats obtenus sur des images de synthèse ainsi que sur des images réelles seront présentés.

En effet, le contenu informationnel des pixels mixtes représente un mélange de plus d'une classe des différentes classes thématiques contenues dans la scène. Ce type de pixels est dû principalement à la limitation de la résolution spatiale et spectrale du capteur utilisé. La limitation de la résolution spatiale du capteur implique que les pixels peuvent couvrir plus d'une classe sur le « terrain » (contexte de "basse résolution")

L'intensité mesurée d'un pixel mixte est alors un mélange d'intensités engendrées par les classes thématiques « pures » sous-jacentes, qui sont appelées des « *endmembers* » figurantes dans la scène.

Il est important de noter que la résolution spatiale et la résolution spectrale ont une relation inextricable où la variabilité spectrale élevée des zones locales de la scène analysée devient apparente quand la résolution spatiale devient plus fine. Ainsi, l'utilisation de capteurs avancés, avec une résolution spatiale supérieure, ne permet pas nécessairement d'enlever la confusion interclasses issue de l'existence des pixels « mixtes ». Par conséquent, l'application des méthodes de démixage des pixels devient très

importante dans le processus de l'interprétation de scène où le contenu sous-pixelique est précieux et des résultats de classification plus précis sont nécessaires.

Une méthode de démixage (en anglais : *unmixing*) vise à identifier les classes pures constituant le spectre mesuré au sein d'un pixel mixte ainsi que la proportion de chacune de ces classes présentes dans le pixel, appelée *taux d'abondance*.

Deux types de modèles sont principalement utilisés par les approches de démixage pour la modélisation de la contribution des différentes classes pures au spectre mesuré à partir d'un pixel mixte : le *modèle linéaire* et les *modèles non-linéaires*. Nous nous contentons, dans ce paragraphe, de décrire le modèle linéaire qui est utilisé dans l'immense majorité des travaux liés au démixage pixelique à cause de sa simplicité de l'application.

Le modèle linéaire est basé sur l'hypothèse selon laquelle l'intensité spectrale mesurée d'un mixel est une combinaison linéaire des intensités spectrales des classes thématiques pures présentes dans ce pixel mixte, pondérées par le taux d'abondance de la classe pure correspondante. Ceci s'exprime mathématiquement de la manière suivante : en supposant X_m la variable aléatoire représentant l'intensité mesurée de la classe thématique C_m , $m = 1, 2, \dots, M$, et τ_m le taux d'abondance de la classe C_m dans un pixel mixte donné, alors, la variable aléatoire X représentant l'intensité mesurée de ce même pixel mixte est donnée par $X = \sum_{m=1}^M \tau_m \cdot X_m$ (avec $\sum_{m=1}^M \tau_m = 1$). Le démixage consiste alors à déterminer les différents taux d'abondance τ_m , $m = 1, 2, \dots, M$ à partir de la seule intensité mesurée X . Il est bien évident que cette opération de démixage est d'une extrême difficulté lorsque l'information disponible sur les pixels mixtes est issue d'un seul capteur. C'est pour cette raison que les travaux qui traitent le démixage considèrent l'utilisation de plusieurs capteurs aboutissant ainsi à plusieurs observations de chaque pixel mixte qui partagent toutes, les mêmes taux d'abondance. A titre d'exemple, en imagerie satellitaire du type multi, ou hyper, spectral, la scène analysée est observée dans N bandes spectrales (Figure 4.10). Dans ce cas, l'intensité spectrale mesurée pour chaque pixel mixte P est un vecteur R_P de dimension N , appelée vecteur de réflectance spectrale. En considérant R la matrice de taille $N \times M$ des réflectances spectrales des classes thématiques « pures » dans chaque bande spectrale, et A le vecteur de taille M comportant les taux d'abondance dans un pixel mixte, nous pouvons alors exprimer la relation entre le vecteur de réflectance spectrale R_P , la matrice A et le vecteur des taux d'abondance par la simple relation matricielle :

$$R_P = R \cdot A \quad (4.12)$$

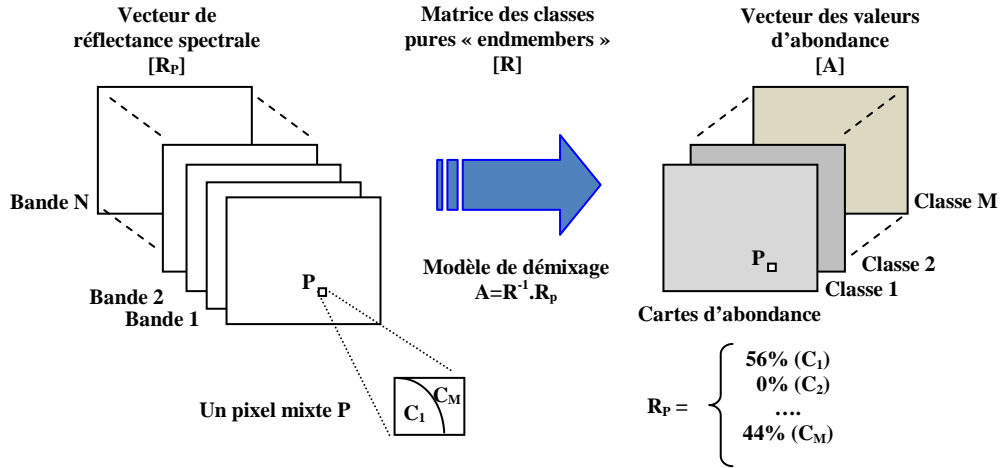


Figure 4.10 : Schéma général d'un processus de démixage spectral utilisant le modèle linéaire

Les méthodes de démixage ont donc pour objectif d'inverser la matrice R afin de pouvoir calculer le vecteur A et établir ainsi des cartes d'abondance (images de même taille que l'image observée) et comportant chacune, le taux d'abondance d'une classe donnée dans tous les pixels de la scène.

Plusieurs méthodes de démixage basées sur ce modèle linéaire ont été proposées dans la littérature. Nous citons les plus utilisées telles que « *PPI : Pixel Purity Index* » [142], « *SMA : Spectral Mixture Analysis* » [39] [143], « *SVD : Singular Value Decomposition* » [144], « *ML : Maximum Likelihood* » [145] et les méthodes basées sur l'analyse en composant indépendantes « *ICA Independent Component Analysis* » [146].

La question essentielle que nous soulevons ici concerne la situation où l'image issue d'un seul capteur est disponible pour réaliser le processus de démixage. Dans ce cas, les méthodes classiques de démixage pixelique se trouvent démunies d'autres informations spectrales, aboutissant ainsi à des résultats médiocres de démixage.

En effet, dans ce contexte que nous appelons mono-capteur, nous proposons de considérer l'hypothèse que chaque pixel observé se trouve dans un contexte spatial qui lui est similaire. Ceci signifie que la composition d'un pixel, en termes de contenu thématique, est quasi identique à la composition des pixels se trouvant dans son voisinage spatial très proche. Sous cette hypothèse, nous proposons de considérer les pixels se trouvant dans le voisinage spatial proche comme étant des réalisations aléatoires du pixel central. Notons que cette proposition consiste à substituer la dimension multi spectrale (ou plus généralement : multi-capteur) par la dimension spatiale.

L'approche de démixage pixelique que nous proposons, consiste à estimer à partir du voisinage local de chaque pixel P de l'image, la distribution de possibilités π_p encodant la classe thématique de ce pixel (même démarche que nous avons adopté pour la classification pixelique), et de considérer les mesures de similarité possibilistes (après normalisation) comme étant les mesures des taux d'abondance correspondant à la composition du pixel considéré [147]:

$$\tau_m = \text{Sim}_{\infty}(\pi_{C_m}, \pi_p) / \sum_{k=1}^M \text{Sim}_{\infty}(\pi_{C_k}, \pi_p), m=1, 2, \dots, M \quad (4.13)$$

Dans la plupart des applications sur des images réelles, il est très difficile de déterminer la composition exacte des pixels mixtes. C'est pourquoi nous proposons dans la suite d'évaluer cette approche sur une image de synthèse pour laquelle nous fixons, *a priori*, les taux d'abondance durant le

processus de génération de cette image. Le modèle linéaire est adopté pour la génération des variables aléatoires représentant les intensités engendrées.

4.6.1 Démixage possibiliste d'images simulées

L'image de test que nous proposons d'utiliser est de taille 550×550 , et comporte deux classes thématiques. Elle est composée de onze secteurs suivants :

- Secteur 1 : secteur comportant des pixels « pures » de la classe C_1 ;
- Secteur 2 : secteur comportant des pixels « pures » de la classe C_2 ;
- Secteur 3 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 10\%$ et $\tau_2 = 90\%$;
- Secteur 4 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 20\%$ et $\tau_2 = 80\%$;
- Secteur 5 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 30\%$ et $\tau_2 = 70\%$;
- Secteur 6 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 40\%$ et $\tau_2 = 60\%$;
- Secteur 7 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 50\%$ et $\tau_2 = 50\%$;
- Secteur 8 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 60\%$ et $\tau_2 = 40\%$;
- Secteur 9 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 70\%$ et $\tau_2 = 30\%$;
- Secteur 10 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 80\%$ et $\tau_2 = 20\%$;
- Secteur 11 : secteur comportant des pixels mixtes avec les taux d'abondance suivants : $\tau_1 = 90\%$ et $\tau_2 = 10\%$;

L'image de synthèse ainsi générée est illustrée dans la Figure 4.11, dans le cas où les deux classes C_1 et C_2 sont supposées suivre des lois de probabilité gaussiennes : $G(m_1=100, \sigma_1=15)$ pour C_1 , et, $G(m_2=150, \sigma_2=15)$ pour C_2 .

Les zones d'apprentissage positionnées par l'expert sont destinées à réaliser l'estimation des distributions de possibilités des deux classes thématiques (en suivant la démarche développée dans le chapitre précédent).



Figure 4.11 : Image de synthèse composée de deux classes C_1 et C_2 , zones d'apprentissage et mélange linéaire de C_1 et C_2

Les résultats obtenus, par l'application de l'approche proposée sur l'image de synthèse des classes gaussiennes de la Figure 4.11, en termes de cartes d'abondance sont données dans la Figure 4.12.

Par simple inspection visuelle de ces cartes d'abondance, on constate la variation linéaire des valeurs d'abondance dans les cartes correspondantes aux deux classes.

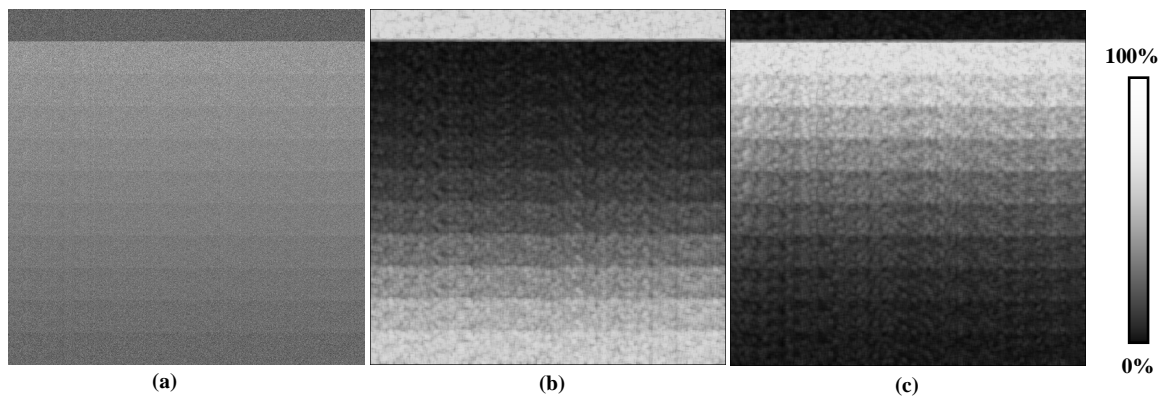


Figure 4.12 : (a) Image de synthèse composée de deux classes générées par deux distributions gaussiennes; (b) carte d'abondance de la classe C_1 et (c) carte d'abondance de la classe C_2

La même conclusion peut être formulée lorsque l'image de synthèse est construite pour deux classes C_1 et C_2 supposées suivre des lois de probabilités du type Rayleigh : $R(\alpha_1 = 30)$ pour C_1 , et, $R(\alpha_2 = 70)$ pour C_2 , (Figure 4.13)

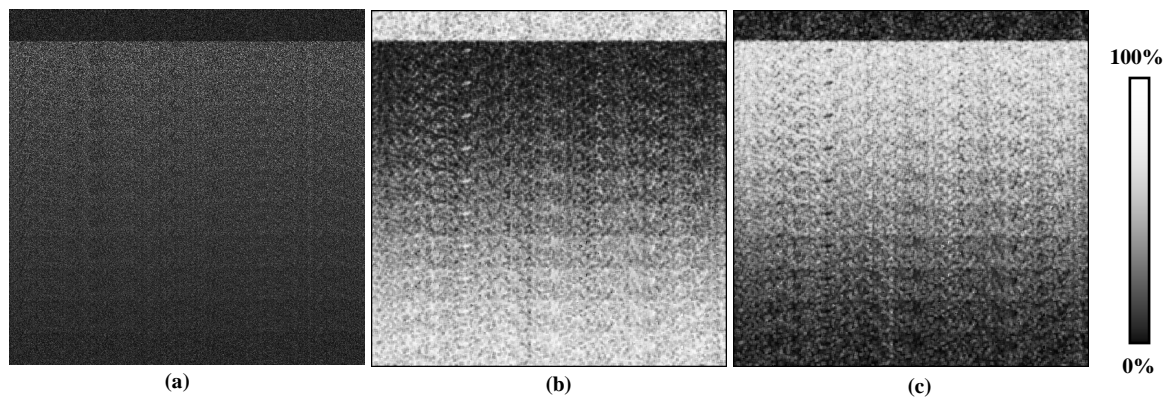


Figure 4.13 : (a) Image de synthèse composée de deux classes générées par deux distributions Rayleigh ; (b) carte d'abondance de C_1 et (c) carte d'abondance de C_2

Une étape d'analyse quantitative des cartes d'abondances est menée afin de confirmer les résultats de l'inspection visuelle. Dans chaque secteur de l'image de synthèse, la moyenne et l'écart type des valeurs d'abondance calculées par l'approche proposée sont données dans le Tableau 1 (cas gaussien) et Tableau 2 (cas Rayleigh) :

Tableau 1 : Valeurs d'abondance des classes C_1 et C_2 (cas gaussien) dans chaque secteur de l'image de synthèse

	Secteur 3	Secteur 4	Secteur 5	Secteur 6	Secteur 7	Secteur 8	Secteur 9	Secteur 10	Secteur 11
	$C_1(10\%)$ $C_2(90\%)$	$C_1(20\%)$ $C_2(80\%)$	$C_1(30\%)$ $C_2(70\%)$	$C_1(40\%)$ $C_2(60\%)$	$C_1(50\%)$ $C_2(50\%)$	$C_1(60\%)$ $C_2(40\%)$	$C_1(70\%)$ $C_2(30\%)$	$C_1(80\%)$ $C_2(20\%)$	$C_1(90\%)$ $C_2(10\%)$
Moyenne (C_1)%	14	20	28	39	50	61	72	79	87
Ecart type (C_1)	0.07	0.08	0.09	0.10	0.10	0.09	0.08	0.08	0.06
Moyenne (C_2)%	86	80	72	61	50	39	28	21	13
Ecart type (C_2)	0.06	0.07	0.09	0.10	0.10	0.10	0.09	0.07	0.06

Tableau 2 : Valeurs d'abondance des classes C_1 et C_2 (cas Rayleigh) dans chaque secteur de l'image de synthèse

	Secteur 3	Secteur 4	Secteur 5	Secteur 6	Secteur 7	Secteur 8	Secteur 9	Secteur 10	Secteur 11
	$C_1(10\%)$ $C_2(90\%)$	$C_1(20\%)$ $C_2(80\%)$	$C_1(30\%)$ $C_2(70\%)$	$C_1(40\%)$ $C_2(60\%)$	$C_1(50\%)$ $C_2(50\%)$	$C_1(60\%)$ $C_2(40\%)$	$C_1(70\%)$ $C_2(30\%)$	$C_1(80\%)$ $C_2(20\%)$	$C_1(90\%)$ $C_2(10\%)$
Moyenne (C_1)%	18	24	32	42	51	62	71	77	81
Ecart type (C_1)	0.1	0.08	0.088	0.06	0.09	0.08	0.1	0.07	0.08
Moyenne (C_2)%	82	76	68	58	49	38	29	23	19
Ecart type (C_2)	0.09	0.08	0.09	0.10	0.09	0.07	0.06	0.08	0.08

Nous remarquons que dans tous les cas de figures, l'erreur moyenne d'estimation des taux d'abondance est inférieure à 10%. Cette erreur augmente lorsque la présence d'une classe thématique dans les pixels mixtes est faible (inférieure à 20%), et diminue lorsque ce taux de présence augmente. Il est aussi très important de noter les faibles valeurs de l'écart-type ce qui montre une bonne compatibilité avec les valeurs des taux d'abondance utilisées pour la génération des secteurs d'images de synthèse.

4.6.2 Démixage possibiliste d'images réelles

L'application de cette approche sur des images réelles montre aussi des résultats dignes d'analyse. En effet, dans certains types d'images, la description des classes thématiques par l'expert est une tâche difficile. Ceci rend le processus d'interprétation plus difficile. Par exemple, dans le cas de la mammographie, la difficulté d'interprétation de ces images est due essentiellement à la variété de la densité des tissus, des structures compliquées du sein, de la grande diversité existante dans les zones de tumeur en termes de type, de forme, de contours, etc. Ainsi, l'expert a souvent besoin d'un système d'aide à l'interprétation réalisant par exemple, une opération de segmentation préliminaire de ces images. Malheureusement, aucun système de segmentation ne peut assurer une segmentation totalement fiable permettant de donner une idée précise sur les zones tumorales [111].

Dans ce paragraphe, nous allons montrer l'efficacité de l'approche proposée afin de mettre en évidence le contenu des régions d'intérêt dans une telle image. Pour ce faire, nous proposons d'appliquer la démarche de démixage pixelique sur des images mammographiques extraites de la base d'images MIAS. L'image donnée dans la Figure 4.14 montre un exemple de la difficulté d'interprétation de ce type d'images. Il s'agit d'images comportant des régions de faible contraste et fortement texturées. L'estimation des distributions de possibilités des deux classes de tissus normaux et tumoraux, à partir de zones caractéristiques fixées par l'expert, montre le grand chevauchement des intensités mesurées entre ces deux classes.

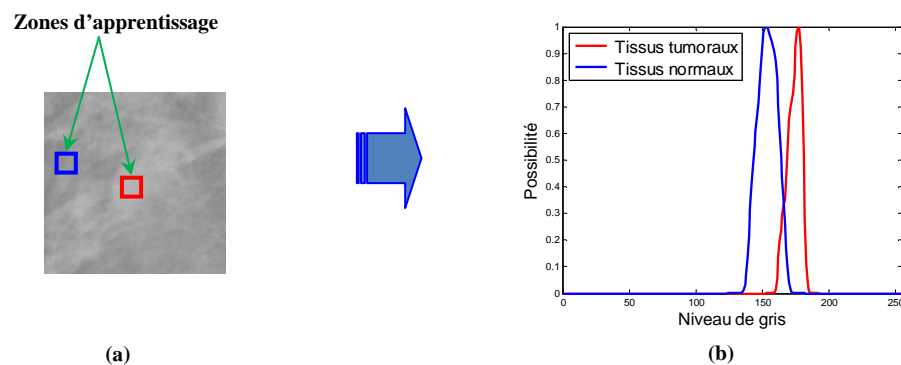


Figure 4.14 : (a) Image mammographique de la base MIAS, (b) distributions de possibilités de tissus normaux et tumoraux estimées à partir des zones caractéristiques fixées par l'expert

Dans un premier temps et afin d'avoir une première idée des régions d'intérêt, une phase de classification pixelique, en utilisant l'approche proposée, a été appliquée sur cette image, Figure 4.15-(a). La segmentation de cette image permet de mettre en avant plusieurs régions d'intérêt de différentes formes et tailles. Malgré le fait que cette segmentation permette à l'expert de focaliser son attention sur

certaines régions, elle ne permet pas d'aller plus loin dans le processus d'interprétation en délivrant plus de détails concernant le contenu sémantique de l'image et notamment sur les foyers tumoraux, par exemple.

La carte d'abondance obtenue par l'approche de démixage proposée et correspondant à la classe du tissu tumoral est donnée dans la Figure 4.15-b et elle est superposée à l'image originale dans la Figure 4.15-c. Cette carte d'abondance représente les mesures de similarité possibilistes après normalisation (les valeurs d'abondance), calculées entre la distribution de possibilité de la classe Tissus tumoraux et chacune des distributions de possibilités locales π_p .

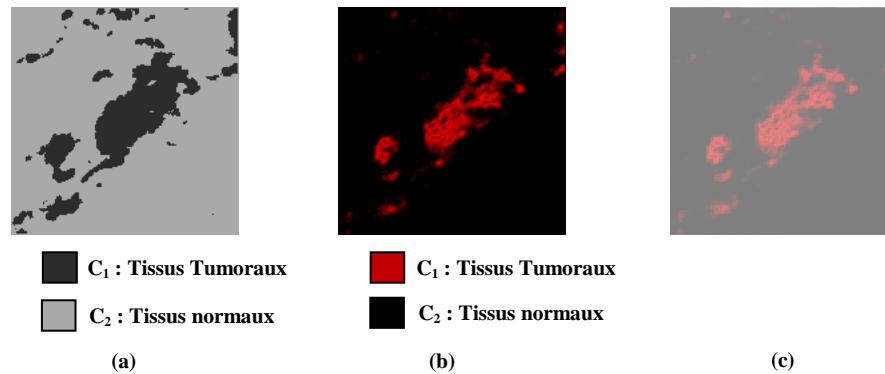


Figure 4.15 : (a) Image mamographique segmentée par l'approche proposée, (b) les valeurs d'abondance de la classe « tumeur » et (c) la superposition des valeurs d'abondance de la classe « tumeur » et l'image originale

L'analyse visuelle des résultats de démixage (i.e. les valeurs d'abondance de tissus tumoraux) permet d'explorer la distribution spatiale des valeurs d'abondance de tissus tumoraux tout en mettant en avant les zones ayant des grandes valeurs d'abondances (foyer tumoraux). Une telle analyse permet à l'expert de constituer une idée plus précise sur le contenu et la localisation des foyers tumoraux (généralement des petites régions d'intérêt) au-delà des foyers primitifs (généralement des grandes régions d'intérêt) ainsi que l'étendue de tissus tumoraux. Ceci constitue des informations précieuses pour le clinicien lui permettant une interprétation plus fine des images analysées.

4.7 Conclusions

Dans ce chapitre, nous avons présenté, dans un premier lieu, le concept de similarité ainsi que les différentes catégories de mesures de similarité proposées dans la littérature. Ensuite, nous avons présenté la notion de la similarité possibiliste et son intérêt dans le processus d'interprétation de scène. Nous avons particulièrement insisté sur l'importance de cette notion pour formaliser la discrimination entre les distributions de possibilités tout en soulignant le problème de sélection d'une mesure de similarité adéquate pour le problème à résoudre. Afin de donner un consensus relatif au choix d'une mesure de similarité, nous avons proposé une méthode d'évaluation des mesures de similarité en classification pixelique. Cette méthode nous a permis de choisir la mesure de similarité qui maximise la discrimination entre distributions de possibilités. En se basant sur le résultat de cette méthode, nous avons développé une approche permettant la classification pixelique sur la base des mesures de similarité possibilistes permettant l'intégration d'informations contextuelles spatiales et des connaissances *a priori*. En effet, l'un des points-clés de l'approche proposée (dans une perspective de classification et de démixage pixelique), est de caractériser chaque pixel à classer en prenant en compte son voisinage au travers de la constitution de distributions de possibilités locales. L'autre point-clé de notre approche est l'utilisation

de la similarité entre distributions de possibilités locales et distributions de possibilités des classes, et non pas entre les degrés de possibilité d'appartenance des paramètres mesurés ponctuellement, aux différentes distributions de possibilités des classes. Les premiers résultats tant sur des images synthétiques (en comparaison avec une approche Markovienne) que sur des images médicales réelles semblent prometteurs. Les étapes suivantes consistent à valider ces premiers résultats sur différents types d'images ayant un nombre de classes supérieur à deux.

Chapitre 5

SYSTEME DE SEGMENTATION/INTERPRETATION POSSIBILISTE D'IMAGE

5.1 Introduction

La segmentation est l'une des opérations de traitement d'image « bas niveau » ayant pour but de regrouper les pixels d'une image connectés par une certaine relation, en des régions homogènes (selon des critères prédéfinis). Il s'agit donc, d'une opération qui permet de passer du niveau des mesures numériques associées aux différents pixels, au niveau de régions ayant un sens sémantique plus élevé que les données en sortie du capteur.

A ce jour, il existe un nombre important de méthodes de segmentation d'image que l'on peut regrouper en trois principales catégories : les méthodes dites orientées *régions*, qui visent à regrouper des pixels connectés ayant des propriétés similaires en des régions homogènes, les méthodes dites orientées *contours*, qui cherchent à détecter les frontières des changements brusques dans l'image (par exemple, la détection des changements dus aux contours des régions ou objets) et les méthodes dites *hybrides*, basées sur la mise en place d'une coopération entre les deux premières méthodes de segmentation [100].

Par ailleurs, en fonction de leur nature itérative, les méthodes de segmentation d'image peuvent également être répertoriées en deux catégories : les méthodes *non itératives* qui réalisent la segmentation en une seule étape et les méthodes *itératives* qui aboutissent à une image segmentée après plusieurs étapes itératives tout en exploitant les résultats intermédiaires de ces étapes [111]. Les méthodes de segmentation d'image par croissance de région « *Region Growing* » [115] constituent l'un des exemples les plus fréquemment rencontrés parmi les méthodes de segmentation itérative. Le principe de ces méthodes consiste en la sélection d'un ensemble de pixels initiaux appelés *germes* qui constituent les

points de départ des régions finales à identifier. Ces méthodes examinent, ensuite, si les pixels voisins des germes initiaux ont des propriétés similaires à celles des germes et déterminent ainsi les pixels voisins que l'on doit ajouter aux différentes régions. Ce processus est répété d'une façon itérative et similaire à celle utilisée dans le regroupement « *Clustering* » (voir chapitre 1 et 3).

Notons que l'immense majorité des méthodes de segmentation d'image opèrent sur les données de bas niveau qui sont les mesures numériques issues des capteurs (i.e. les intensités associées aux différents pixels) ou sur des grandeurs numériques calculées à partir de ces mesures. Cette forte dépendance, par rapport aux capteurs, constitue un inconvénient majeur de ces méthodes de segmentation car les régions détectées sont homogènes selon des critères appliqués aux données issues des capteurs et non selon des critères basés sur le contenu sémantique des régions constituant la scène observée et que nous cherchons à interpréter. Ceci explique l'intérêt croissant pour l'intégration au sein des méthodes de segmentation d'image des connaissances *a priori* liées au contenu sémantique des images à segmenter.

Notons que suite à la segmentation d'image en des régions homogènes, les systèmes d'interprétation de scène appliquent une opération de classification d'image qui, sur la base de connaissances sémantiques du haut niveau (liées aux potentielles classes thématiques présentes dans la scène analysée) consiste à attribuer un label de classe à chaque région homogène détectée.

Cette intégration des connaissances « haut niveau » pour un traitement collaboratif avec les données bas niveau occupe une place très importante dans les activités de recherche actuelles en vision par ordinateur. A ce stade, deux questions sont soulevées : quelles sont les connaissances de type haut niveau que l'on considère intéressantes à intégrer (et comment peut-on les représenter ?), et quels sont les outils permettant d'exploiter ces connaissances en segmentation d'image et interprétation de scène ?

Parmi ces connaissances, nous nous focalisons sur les connaissances exprimées par l'expert d'une manière vague via des contraintes épistémiques réalisant ainsi, la projection de ses connaissances sémantiques sur les propriétés des quantités physiques mesurées des régions observées sur l'image analysée. Ceci justifie notre choix de l'application du cadre possibiliste pour l'analyse et la segmentation des images considérées.

A l'issue de la segmentation en régions homogènes, la phase de classification d'image consiste à analyser chaque région homogène identifiée afin de lui attribuer un sens sémantique comme étant :

- d'une seule classe parmi les classes thématiques prédéfinies (i.e. classification sous l'hypothèse du monde fermé), ou
- une région homogène comportant des mixels qui sont un mélange des classes prédéfinies, ou bien
- une région comportant une nouvelle classe thématique non identifiée initialement par l'expert (i.e. classification sous l'hypothèse du monde ouvert).

La majorité des travaux sur la classification d'image se focalise sur l'hypothèse du monde fermé et se contente de l'attribution d'un label de classe parmi les classes thématiques prédéfinies à chaque région homogène détectée. C'est pour cette raison que beaucoup de travaux en traitement d'image ne font pas la distinction et considèrent comme étant identiques les deux opérations de segmentation et de classification d'image. A ce titre, la théorie des ensembles flous (et son extension vers la théorie des possibilités) apporte un outil fort intéressant pour la segmentation/classification d'image permettant non seulement la fidèle représentation des connaissances ambiguës exprimées par l'expert, mais aussi de considérer l'appartenance partielle d'une région à une classe (incluant aussi la totale appartenance et la totale non appartenance). En effet, l'application de la théorie des ensembles flous en segmentation/classification d'image consiste à transcrire la description des experts sous la forme d'ensembles flous [84] en transformant les mesures numériques issues des capteurs en des valeurs d'appartenance, aboutissant ainsi

à des *images dites floues* (i.e. des images comportant l'appartenance partielle des pixels aux différentes classes). Ainsi, les méthodes de segmentation/classification floues opèrent sur les images floues dans l'objectif d'identifier les régions homogènes et de leur attribuer des appartenances partielles aux différentes classes thématiques. Parmi les méthodes de segmentation/classification floues des images, nous pouvons citer: l'algorithme C-moyennes flous (Fuzzy C-Means, FCM) [94], la classification hiérarchique floue [116] et la classification floue contextuelle itérative [38] (plus de détails concernant ces méthodes de segmentation floue peuvent être trouvés dans [111]).

La théorie de possibilités est aussi appliquée à la question de segmentation/classification d'images [37][72][138][139][140][111]. D'une façon générale, cette application est très similaire à celle utilisée dans le cadre de la théorie des ensembles flous. Elle consiste à traduire la description donnée par l'expert, en des distributions de possibilités décrivant la relation entre les mesures numériques issues des capteurs et les différentes classes thématiques qui composent l'image à segmenter. Ces distributions permettent ainsi de produire des images dites *cartes possibilistes* (correspondant au concept d'images floues) où des opérations de traitement d'image ainsi que des opérations de fusion possibiliste peuvent être appliquées sur ces images.

D'une manière générale, ces méthodes de segmentation/classification d'image suivent une démarche « ascendante » du niveau pixelique vers le niveau objet en caractérisant les classes thématiques par l'intermédiaire des connaissances épistémiques exprimées par l'expert. Cette caractérisation devient très difficile dans un contexte où ces classes sont difficiles à discriminer (faible contraste, environnement très bruité ou existence de classes multimodales). De plus, et malgré le fait que ces méthodes soient basées sur le concept d'appartenance partielle, il est étonnant de constater qu'elles n'ont pas été appliquées pour l'interprétation de la composition des régions homogènes en termes de démixage et d'extraction de connaissances sur la composition/contenu de telles régions en termes de classes constituantes.

Dans ce chapitre, nous proposons une approche de segmentation/interprétation d'image basée sur une démarche possibiliste opérant dans les deux directions : ascendante et descendante (Figure 5.1).

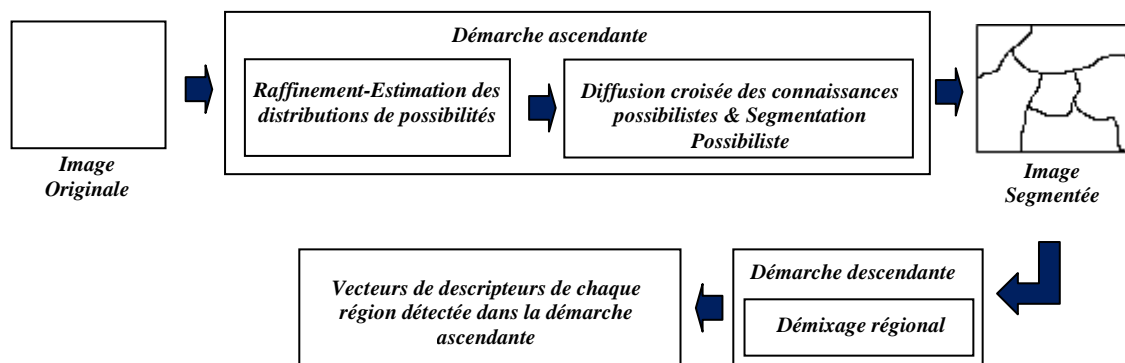


Figure 5.1 : Schéma global de l'approche proposée de segmentation/interprétation d'image basée sur une démarche possibiliste opérant dans les deux directions : ascendante et descendante

Dans la direction ascendante, nous proposons de mener une démarche de segmentation possibiliste d'images composée de deux étapes consécutives : le *raffinement* (visant à construire une base d'apprentissage représentative des différentes classes) et la *diffusion des connaissances possibilistes* [111] (permettant de réduire l'incertitude liée à la classification pixelique). Cette exploitation conjointe des sources de connaissances permet d'accumuler les évidences sur les régions (ou objets) et aboutit à une image segmentée-classifiée de la scène analysée.

La direction descendante que nous proposons concerne une démarche de démixage appliquée au niveau des régions en exploitant le contexte pixelique spatial. Cette démarche permet à l'expert de mettre en cause le contenu informationnel des régions/objets détectés dans la phase ascendante de l'approche proposée. A l'issue de cette démarche descendante, chaque région détectée dans la démarche ascendante peut être caractérisée par un vecteur de descripteurs haut niveau qui ne se limitent pas à un simple étiquetage par un label de classe. En effet, parmi les composants de ce vecteur nous proposons d'intégrer les histogrammes des taux d'abondance de chacune des classes thématiques dans les pixels composants chaque région considérée.

D'un premier lieu, nous rappelons le principe de raffinement des distributions de possibilités. Ensuite, nous présentons le principe de diffusion des connaissances possibilistes et les méthodes utilisées pour effectuer cette diffusion des connaissances. Nous enchainons par l'étude de l'effet d'enrichissement des connaissances d'expert par de nouveaux échantillons sur le processus de segmentation dans un contexte où les classes sont difficiles à discriminer. Finalement, nous présenterons la démarche descendante permettant de déterminer, pour les régions détectées dans la phase ascendante, leurs compositions globales en termes des classes thématiques de base.

5.2 Analyse ascendante pour la segmentation possibiliste

L'interprétation humaine du contenu d'une scène consiste à focaliser l'attention, d'abord sur les objets les plus apparents dans la scène et qui peuvent être considérés comme des « références » pour l'interprétation. Les connaissances liées à ces objets identifiés dans la scène vont ensuite permettre de guider plus finement l'interprétation progressive en émettant des hypothèses précises sur le voisinage des objets de référence [7]. Ce raisonnement consiste donc, à réaliser une propagation par diffusion des connaissances visant à valider l'existence des nouveaux objets.

Une première modélisation de ce type de raisonnement pour la segmentation d'image en exploitant l'information issue du contexte spatial des objets a été proposée dans [111]. Dans cette modélisation possibiliste, la caractérisation des différentes classes thématiques présentes dans l'image est directement exprimée par l'expert sous la forme de distributions de possibilités (ayant des formes prédéfinies). Cette caractérisation, injectée par l'expert, est ensuite exploitée afin de construire des cartes possibilistes correspondant aux différentes classes présentes dans la scène. L'analyse de ces cartes possibilistes, a permis d'identifier les zones les plus caractéristiques des classes désignées par l'expert. Etant donnée la faible connaissance injectée par l'expert, ces zones sont souvent réduite à des pixels isolés pour lesquels nous avons une forte confiance d'appartenir aux différentes classes thématiques. Ces pixels sont appelés des germes possibilistes. En se basant sur ces germes, un processus de diffusion des connaissances possibilistes est réalisé via un raisonnement possibiliste appliqué aux informations contextuelles spatiales et ceci dans le but de classifier des nouveaux pixels.

Le choix par l'expert d'une forme discriminante de distributions de possibilités pour la caractérisation des différentes classes devient très difficile dans le contexte où ces classes sont relativement difficiles à discriminer (faible contraste, environnement très bruité ou existence de classes multimodales). Par conséquence, les objets de référence représentant les classes désignées par l'expert, sont réduits à des « pixels isolés » de référence.

Afin de remédier à ce problème présent dans le système de segmentation d'image proposé dans [111], nous proposons une démarche de segmentation possibiliste en deux phases. La première phase consiste à réaliser le processus d'estimation-raffinement précédemment détaillé, des distributions de possibilités exprimées par l'expert dans l'objectif d'extraire de véritables objets de référence caractérisant les

différentes classes. La deuxième phase consiste alors en l'application du processus de diffusion des connaissances possibilistes spatiales en se basant sur les objets-germes extraits dans la première phase.

5.2.1 Raffinement-Estimation des distributions de possibilités

Dans ce travail, nous avons proposé une approche itérative de raffinement des distributions de possibilités. Sur la base d'une distribution de possibilités initiale (exprimée par un expert soit sous la forme d'une forme standard ou en sélectionnant un ensemble réduit d'échantillons), l'approche proposée consiste à appliquer un raisonnement possibiliste spatial dans l'objectif d'enrichir une base d'apprentissage par des échantillons, i.e. des pixels, caractéristiques.

L'application itérative de l'approche proposée (dans les deux contextes de faible et de grande complexité de discrimination entre les classes thématiques présentes dans la scène analysée) a permis d'aboutir à des distributions ayant pour caractéristique de se focaliser sur les germes les plus représentatifs de chaque classe thématique, et de « négliger » l'ensemble des échantillons se trouvant dans la queue des distributions réelles représentant ces mêmes classes.

Dans l'approche de segmentation possibiliste que nous proposons, l'application du processus itératif de raffinement-estimation vise l'extraction de la majorité des germes caractéristiques. Ensuite, nous proposons de considérer la projection possibiliste de ces germes, sous la forme de cartes possibilistes correspondantes aux différentes classes thématiques, comme étant les *objets possibilistes de référence* pour la formation des régions finales dans l'image segmentée. En d'autres termes, notre objectif est d'exploiter les cartes possibilistes issues du processus itératif de raffinement-estimation comme point de départ du processus de segmentation finale en réalisant la diffusion des connaissances possibilistes engendrées par les objets possibilistes de référence ainsi obtenus. Cette diffusion des connaissances possibilistes sera présentée dans la suite.

Pour la suite, nous considérons la scène analysée (et observée via une image I) comme étant formée de M classes thématiques $\psi = \{C_1, C_2, \dots, C_M\}$ prédéfinies par l'expert. A l'issue du processus de raffinement-estimation, M distributions de possibilités raffinées sont obtenues. L'application de ces distributions de possibilités sur l'image I conduit à construire M cartes possibilistes CP_{I,C_m} , $m = 1, \dots, M$.

5.2.2 Diffusion des connaissances possibilistes

Dans le système de segmentation d'image proposé dans [111], deux méthodes de diffusion des connaissances possibilistes ont été proposées : la diffusion possibiliste dite classe/classe et la diffusion possibiliste dite croisée.

La méthode de diffusion possibiliste classe/classe consiste à extraire les régions correspondant aux différentes classes thématiques prédéfinies en réalisant une mise-à-jour des M cartes possibilistes d'une façon indépendante les unes des autres. Dans cette méthode, une fonction « g » (Equation (5.1)), issue du domaine du traitement d'image (par exemple, réalisant un filtrage moyen), est exploitée afin de diffuser les connaissances possibilistes dans le contexte spatial des différents pixels (Figure 5.2).

$$CP_{I,C_m}(t+1) = g(CP_{I,C_m}(t)), \quad m=1, \dots, M \quad (5.1)$$

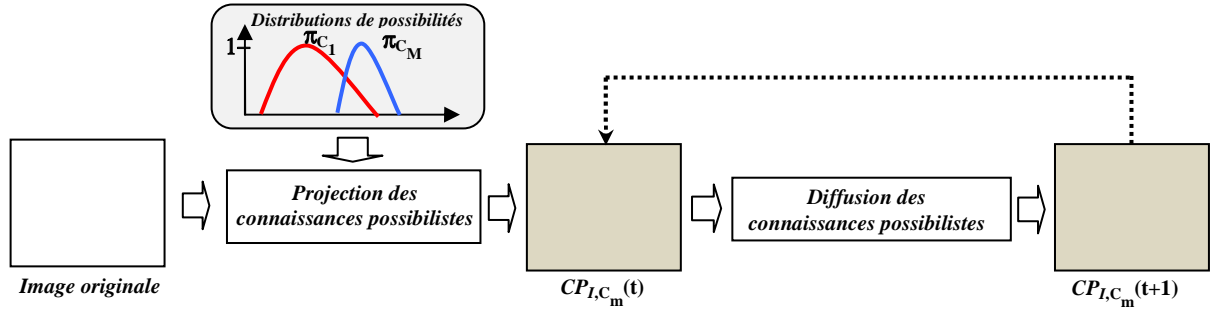


Figure 5.2 : Méthode de diffusion des connaissances possibilistes classe/classe pour une classe thématique donnée

En revanche, la méthode de diffusion possibiliste croisée consiste à exploiter l'ensemble des M cartes possibilistes pour réaliser la mise à jour de chacune d'entre elles d'une façon itérative. Dans cette méthode, une fonction « Q » est exploitée afin de diffuser les connaissances possibilistes dans le contexte spatial des différents pixels (Figure 5.3)

$$CP_{I,C_m}(t+1) = Q(CP_{I,C_1}(t), \dots, CP_{I,C_m}(t), \dots, CP_{I,C_M}(t)), \quad m=1, \dots, M \quad (5.2)$$

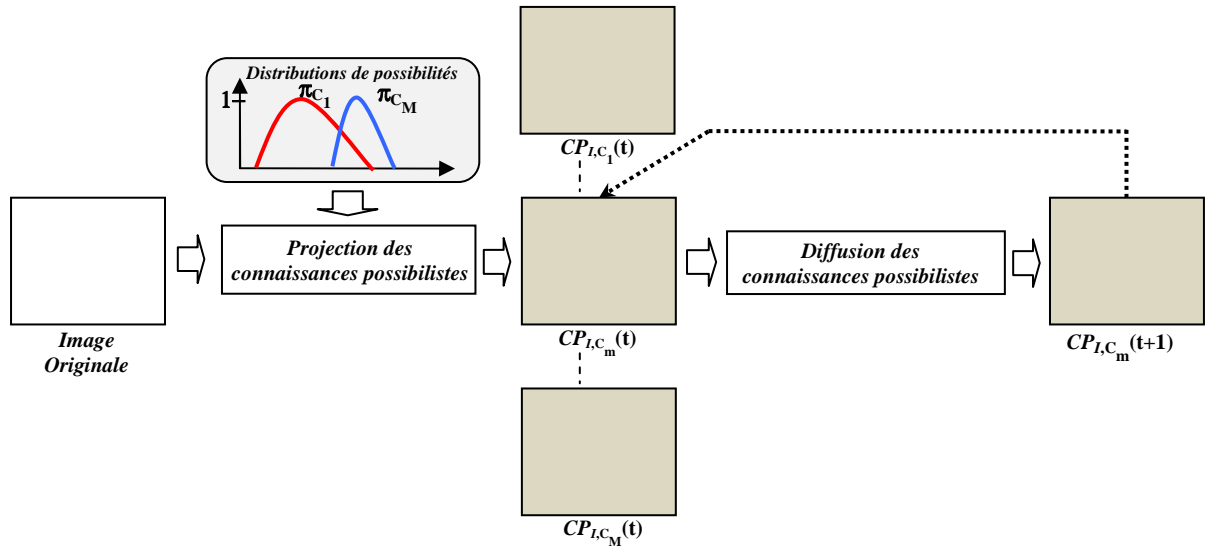


Figure 5.3 : Méthode de diffusion croisée des connaissances possibilistes

La fonction adoptée pour réaliser la diffusion des connaissances possibilistes dans les deux méthodes est basée sur l'utilisation des sous-fenêtres définies par le filtre de Nagao [117]. Le choix de ce filtre est dû à sa capacité à conserver les caractéristiques des pixels proches des contours.

En effet, notons $F = \{f_1, f_2, \dots, f_9\}$, l'ensemble des 9 sous-fenêtres spatiales, appelées *cliques*, définies par le filtre de Nagao et contenues dans le voisinage de chaque pixel (par exemple, la configuration présentée dans la Figure 5.4 pour un voisinage 5×5 de chaque pixel).

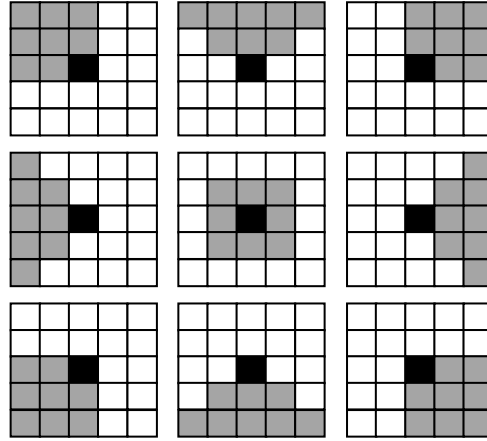


Figure 5.4 : Configuration des sous-fenêtres (ou cliques) du filtre de Nagao

Pour un pixel P donné, chaque carte possibiliste CP_{I,C_m} , $m = 1, 2, \dots, M$, apporte 9 informations possibilistes spatiales correspondant aux neuf cliques. Chacune de ces informations spatiales contient la valeur moyenne des valeurs de possibilité véhiculées par la clique considérée au niveau de la carte possibiliste donnée. Cette représentation permet de résumer l'ensemble des connaissances possibilistes sous la forme d'une matrice possibiliste ayant 9 lignes (correspondante chacune à une clique) et M colonnes (correspondant chacune à une classe donnée), comme illustré dans la Figure 5.5.

Ainsi, chaque clique f_n permet de définir, à une itération « t » donnée, une distribution de possibilités $\pi_{t,n}$ sur l'ensemble des M classes thématiques $\psi = \{C_1, C_2, \dots, C_M\}$ où

$$\pi_{t,n}(C_m) = \frac{1}{\text{Card}(f_n)} \sum_{\{i,j\} \in f_n} \pi_{t,(i,j)}(C_m) \quad (5.3)$$

La valeur $\pi_{t,n}(C_m)$ représente dans quelle mesure il est possible, à l'itération « t », que la classe C_m soit l'unique singleton qui s'est produite étant donné le clique f_n .

Pour chaque pixel considéré P , la méthode de diffusion classe/classe des connaissances possibilistes consiste à sélectionner une sous-fenêtre gagnante au niveau de chaque carte possibiliste, et ceci afin de mettre à jour les connaissances à l'itération suivante.

$$\pi_{t+1,P}(C_m) = \max(\pi_{t,n}(C_m)), n=1, \dots, 9 \quad (5.4)$$

En revanche, la méthode de diffusion croisée des connaissances possibilistes consiste à identifier (pour chaque pixel) une sous-fenêtre « optimale » parmi l'ensemble des sous-fenêtres de Nagao associées à l'ensemble des classes et d'utiliser cette même sous-fenêtre afin de mettre à jour l'ensemble des cartes possibilistes.

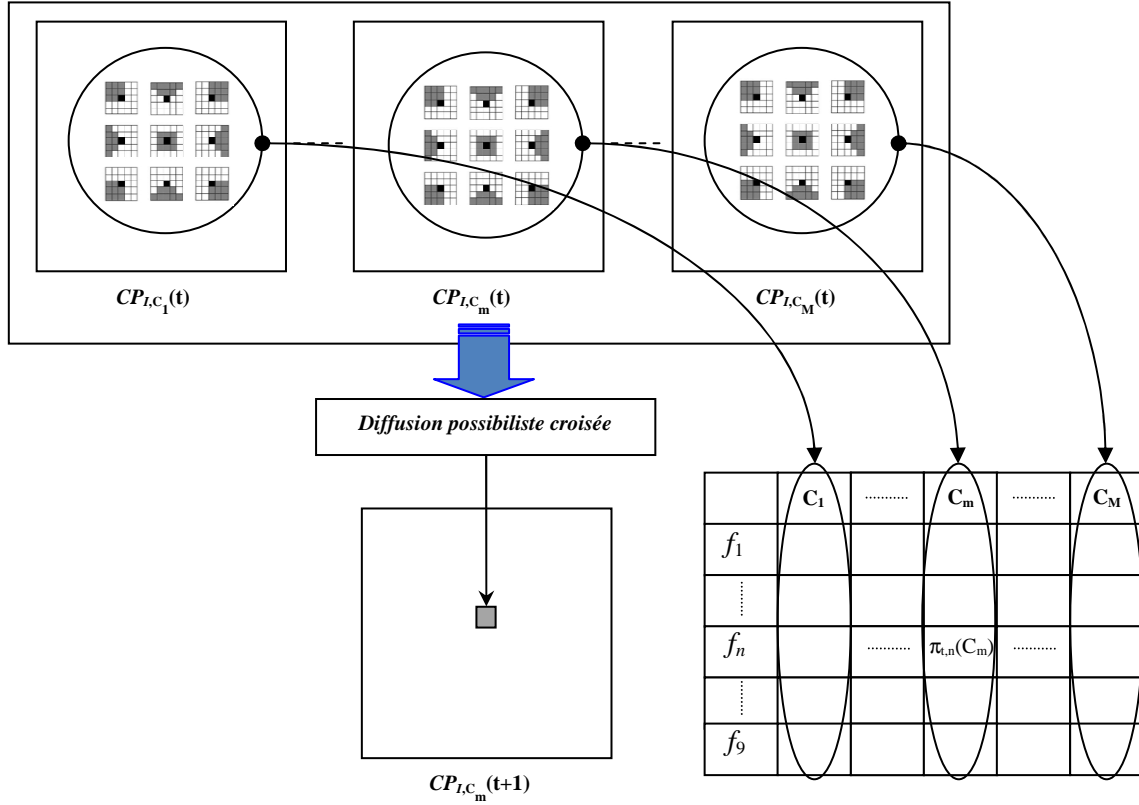


Figure 5.5 : Méthode de diffusion croisée de diffusion des connaissances possibilistes basée sur les cliques de Nagao

Dans cette méthode, l'indice de confiance est utilisé comme mesure d'incertitude permettant d'évaluer la quantité d'information utile pour la prise de décision dans chacune des neuf distributions de possibilités, correspondant aux neuf cliques de Nagao. Alors, à une itération t donnée, et en considérant une clique f_n , $n = 1, 2, \dots, 9$, chaque classe thématique sera caractérisée par la mesure de l'indice de confiance :

$$\begin{aligned}
 Ind_{t,n}(C_m) &= \Pi(\{C_m\}) + N(\{C_m\}) - 1 \\
 &= \pi_{t,n}(C_m) + 1 - \Pi(\{C_m\}^c) - 1 \\
 &= \pi_{t,n}(C_m) + 1 - \max_{k \neq m} \pi_{t,n}(C_k) - 1 \\
 &= \pi_{t,n}(C_m) - \max_{k \neq m} \pi_{t,n}(C_k)
 \end{aligned}$$

Ainsi, un indice de confiance global associé à chaque clique f_n , $n = 1, 2, \dots, 9$, est défini par :

$$Ind_{t,n} = \max_m Ind_{t,n}(C_m) \quad (5.4)$$

La clique optimale, qui est considérée pour la mise à jour des connaissances possibilistes, est celle ayant l'indice de confiance maximal

$$f_{t,n}(\text{optimale}) \longleftrightarrow Ind_{t,n}(\text{optimale}) = \max_{n \in \{1, 2, \dots, 9\}} Ind_{t,n}$$

ainsi :

$$\pi_{t+1,P}(C_m) = \frac{1}{\text{Card}(f_{t,n}(\text{optimale}))} \sum_{\{i,j\} \in f_{t,n}(\text{optimale})} \pi_{t,(i,j)}(C_m) \quad (5.5)$$

Dans [111] [112], l'évaluation de ces deux approches de diffusion des connaissances possibilistes a été menée sur la base d'images de synthèse ainsi que des images mammographiques. L'approche de propagation par diffusion croisée a montré une harmonisation des connaissances possibilistes et pour la classification pixelique dans les zones homogènes. En revanche, la diffusion du type classe/classe présente le seul avantage d'une harmonisation des connaissances possibilistes dans les zones de contours de séparation entre les différentes classes. Par conséquent, et comme notre objectif concerne la formation des régions homogènes (plutôt que la préservation des contours), nous allons adopter l'approche de diffusion croisée des connaissances possibilistes pour la suite de notre travail.

5.3 Comportement de la démarche de segmentation proposée

Dans ce paragraphe, nous allons évaluer le comportement de l'approche ascendante proposée pour la segmentation possibiliste. Le schéma général de cette approche est donné dans Figure 5.6.

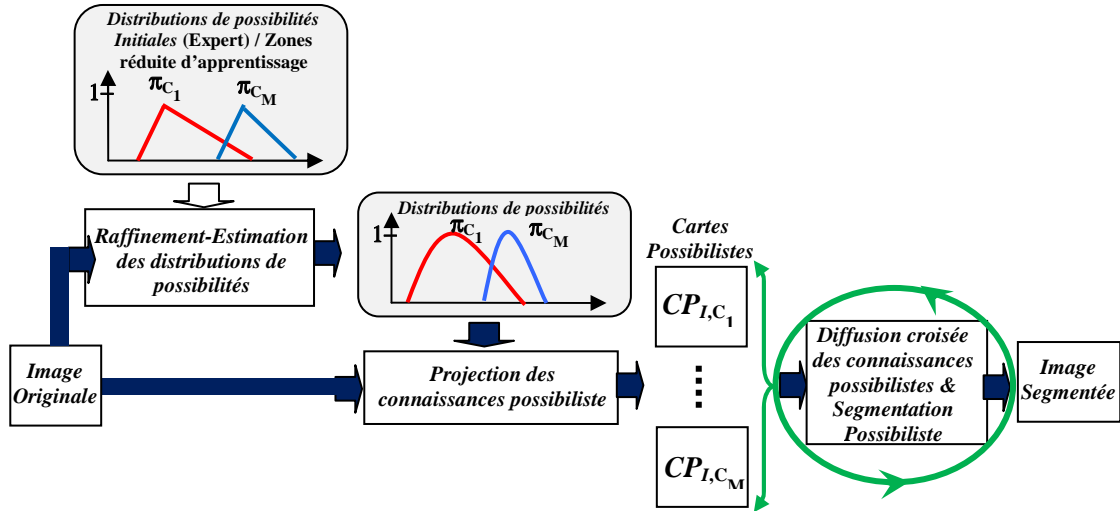


Figure 5.6 : Schéma global de l'analyse ascendante pour la segmentation possibiliste

Cette évaluation est réalisée sur des images de synthèse ainsi que sur des images réelles. Deux types de classes d'objets sont principalement considérés : les classes monomodales (où ces classes sont faciles ou difficiles à discriminer) et le cas des classes multimodales.

5.3.1 Evaluation des performances dans le cas de classes monomodales

Pour l'évaluation des performances de l'approche ascendante de segmentation possibiliste dans le cas de classes monomodales d'objets, nous considérons l'image de synthèse illustrée dans la Figure 3-(c) comportant des objets de différentes tailles d'une classes donnée (C_1) et se trouvant dans un contexte du fond appartenant à une autre classe (C_2). Les pixels des deux classes C_1 et C_2 sont générés par des lois gaussiennes $G(m_1, \sigma_1)$ et $G(m_2, \sigma_2)$ ou par des lois de Rayleigh $R(\alpha_1)$ et $R(\alpha_2)$.

La Figure 5.7 illustre les deux cas considérés :

- Le cas gaussien de faible complexité de discrimination entre les classes (avec les paramètres suivants $\{C_1(160,10), C_2(100,15)\}$) et le cas gaussien de grande complexité de discrimination entre les classes (avec les paramètres suivants $\{C_1(120,15), C_2(100,20)\}$), et

- Le cas de Rayleigh de faible complexité de discrimination entre les classes (avec les paramètres suivants $\{\alpha_1(100), \alpha_2(30)\}$), et le cas Rayleigh de grande complexité de discrimination entre les classes (avec les paramètres suivants $\{\alpha_1(65), \alpha_2(30)\}$).

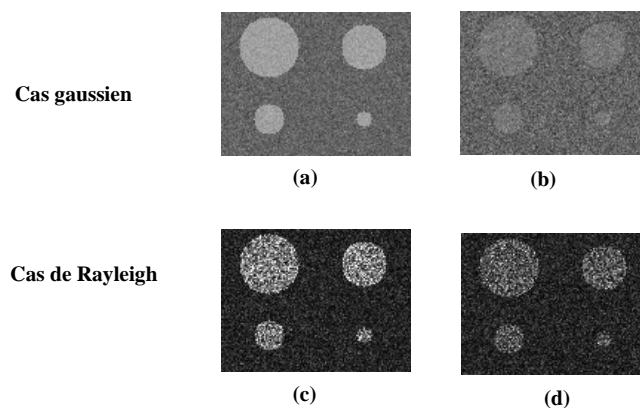


Figure 5.7 : Images de synthèse (cas gaussien) (a) et (b) faible et grande complexité de discrimination entre les classes respectivement et (cas de Rayleigh) (c) et (d) faible et grande complexité de discrimination entre les classes respectivement

Les connaissances initiales d'expert ont été injectées de deux façons : soit en délimitant des zones limitées caractéristiques des classes, soit par une description (en utilisant des variables linguistiques) de certaines propriétés ambiguës caractérisant ces classes. Etant donnée la similarité des résultats obtenus, nous optons pour le choix le plus simple, celui de délimiter par l'expert, des zones caractéristiques de faibles tailles.

La Figure 5.8 illustre les cartes possibilistes obtenues à l'issue du processus de diffusion croisée des connaissances possibilistes et ceci pour les deux cas gaussien et de Rayleigh en faible et grande complexité de discrimination entre les classes. Une première analyse visuelle de ces cartes montre l'homogénéité importante des régions détectées pour les deux classes C_1 et C_2 . L'analyse visuelle des résultats en termes d'images segmentées, données sur la même figure, montre la bonne détection des objets recherchés, et ceci pour les différentes tailles d'objets considérés. Le bruit du type « sel et poivre » sur les images segmentées augmente en fonction de la complexité des images considérées mais n'altère pas la bonne détection des objets d'intérêt majeur.

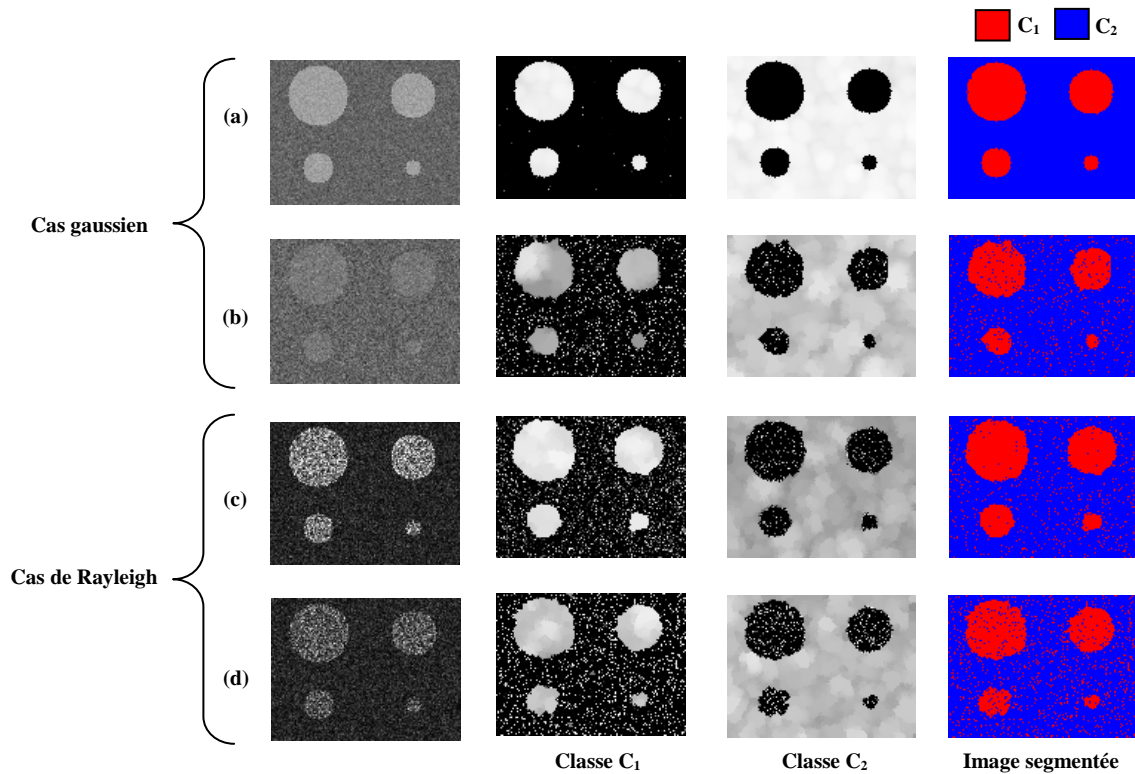


Figure 5.8 : Cartes possibilistes des classes C_1 et C_2 et les résultats de la segmentation à l'issue de la démarche de segmentation proposée (a) et (b) dans le cas gaussien avec une faible et grande complexité de discrimination entre les classes respectivement et (c) et (d) dans le cas de Rayleigh avec une faible et grande complexité de discrimination entre les classes respectivement.

Etant donné que nous appliquons une approche itérative de segmentation possibiliste, la Figure 5.9 représente l'évolution, en fonction du nombre d'itérations, de l'erreur quadratique moyenne entre les cartes possibilistes successives obtenues au cours de la diffusion croisée des connaissances possibilistes et ceci dans les deux cas considérés. Nous constatons la grande rapidité de la convergence du processus de diffusion croisée des connaissances possibilistes. En effet, l'erreur quadratique moyenne (et ceci pour les deux classes C_1 et C_2) converge vers la valeur zéro en moins de douze itérations dans les deux cas : gaussien et Rayleigh.

Une étape de classification pixelique basée sur la règle de décision de la possibilité maximale a été appliquée sur l'image de synthèse dans les deux cas considérés : gaussien et Rayleigh et dans les quatre configurations suivantes :

- **avant** raffinement des distributions de possibilités et **sans** diffusion des connaissances (i.e. seules les connaissances de l'expert sont exploitées, les connaissances contextuelles ne sont pas exploitées) ;
- **avant** raffinement des distributions de possibilités et **avec** la diffusion croisée des connaissances possibilistes (i.e. seules les connaissances de l'expert sont exploitées pour la diffusion des connaissances possibilistes) ;
- **après** raffinement des distributions de possibilités et **sans** diffusion des connaissances (i.e. en se basant sur les connaissances spatiales pour l'extraction des germes possibilistes et en affinant l'estimation des distributions de possibilités, aucune diffusion de connaissance n'est opérée) ;
- **après** raffinement des distributions de possibilités et **avec** la diffusion croisée des connaissances possibilistes (i.e. en se basant sur les connaissances spatiales pour l'extraction des germes

possibilistes et en affinant l'estimation des distributions de possibilités, puis en opérant la diffusion des connaissances possibilistes).

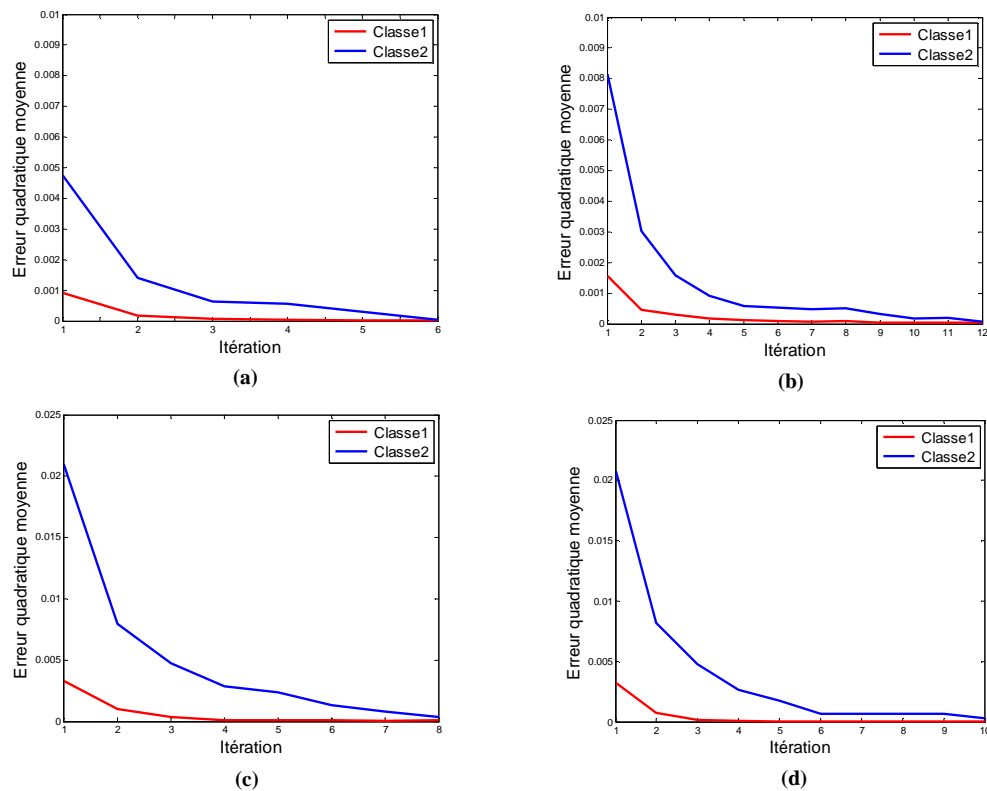


Figure 5.9 : Erreur quadratique moyenne entre deux cartes possibilistes successives dans la diffusion des connaissances possibilistes, (cas gaussien)
 (a) et (b) faible et grande complexité de discrimination entre les classes respectivement et (cas de Rayleigh)
 (c) et (d) faible et grande complexité de discrimination entre les classes respectivement

L'image de synthèse a été générée 100 fois et la moyenne et l'écart-type des taux de reconnaissance sont utilisés comme critères d'évaluation du processus de segmentation possibiliste descendante. Le Tableau 3 (resp. Tableau 4) présente les taux de reconnaissance dans le cas gaussien et pour un faible (resp. grande) complexité de discrimination entre les classes.

Tableau 3. Taux de reconnaissance dans le cas gaussien (et Faible complexité de discrimination) pour les quatre configurations considérées (avec et sans raffinement des distributions de possibilités) et (avec et sans diffusion des connaissances possibilistes)

	Taux de reconnaissance	Ecart type
Avant raffinement et sans diffusion des connaissances	88.19	0.40
Avant raffinement et avec diffusion des connaissances	94.49	1.30
Après raffinement et sans diffusion des connaissances	89.21	2.41
Après raffinement et avec diffusion des connaissances)	<u>97.10</u>	<u>0,09</u>

Comme nous pouvons le constater :

- la diffusion croisée des connaissances possibilistes a permis une nette amélioration du taux de reconnaissance de l'ordre de 6% dans le cas où les connaissances initiales de l'expert (i.e. distributions de possibilités fixées par l'expert) sont utilisées sans raffinement, et de l'ordre de 8% dans le cas où ces distributions de possibilités ont été raffinées) ;
- l'application du seul processus de raffinement des distributions de possibilités n'a apporté qu'une amélioration de l'ordre de 1% en termes de taux de reconnaissance ;
- l'application de la seule diffusion croisée des connaissances a permis une augmentation du taux de reconnaissance d'environ 3% (entre les cas sans et avec raffinement des distributions de possibilités) ; et finalement
- l'association du processus de raffinement des distributions de possibilités avec la diffusion croisée des connaissances possibilistes a permis une amélioration de l'ordre de 9%.

Il est aussi très important de noter la baisse de valeur de l'écart-type ce qui montre une sorte de stabilité des valeurs de taux de reconnaissance ainsi obtenues.

Ces résultats se confirment pour le cas gaussien avec une grande complexité de discrimination entre les classes (Tableau 4). Seul le taux de reconnaissance après raffinement des connaissances de l'expert connaît une diminution de l'ordre de 4% par rapport au cas sans raffinement de ces distributions de possibilités. Ce phénomène a été signalé lors de la présentation du processus d'affinement des distributions de possibilités afin de montrer que l'objectif du processus d'affinement consiste à capturer les germes possibilistes les plus fréquents et les plus caractéristiques, et non d'améliorer le taux de reconnaissance. En revanche, la diffusion croisée des connaissances possibilistes sur la base de l'ensemble des germes capturés par le processus de raffinement des distributions de possibilités a permis une très forte amélioration du taux de reconnaissance (de l'ordre de 17%). Ce qui montre le grand intérêt de l'association entre les deux processus du raffinement et de diffusion des connaissances possibilistes.

Tableau 4 : Taux de reconnaissance dans le cas gaussien (et Grande complexité de discrimination) pour les quatre configurations considérées (avec et sans raffinement des distributions de possibilités) et (avec et sans diffusion des connaissances possibilistes)

	Taux de reconnaissance	Ecart type
Avant raffinement et sans diffusion des connaissances	76.79	2.12
Avant raffinement et avec diffusion des connaissances	80.35	2.58
Après raffinement et sans diffusion des connaissances	72.14	3.85
Après raffinement et avec diffusion des connaissances)	<u>89.26</u>	<u>1.40</u>

De la même façon, le Tableau 5 (resp. Tableau 6) présente le taux de reconnaissance dans le cas de Rayleigh pour le cas de faible (resp. grande) complexité de discrimination entre les classes.

Tableau 5 : Taux de reconnaissance dans le cas de Rayleigh (et Faible complexité de discrimination) pour les quatre configurations considérées (avec et sans raffinement des distributions de possibilités) et (avec et sans diffusion des connaissances possibilistes)

	Taux de reconnaissance	Ecart type
Avant raffinement et sans diffusion des connaissances	80.17	2.15
Avant raffinement et avec diffusion des connaissances	93.34	2.32
Après raffinement et sans diffusion des connaissances	67.5	1.74
Après raffinement et avec diffusion des connaissances)	<u>93.94</u>	<u>1,27</u>

Tableau 6 : Taux de reconnaissance dans le cas de Rayleigh (et Grande complexité de discrimination) pour les quatre configurations considérées (avec et sans raffinement des distributions de possibilités) et (avec et sans diffusion des connaissances possibilistes)

	Taux de reconnaissance	Ecart type
Avant raffinement et sans diffusion des connaissances	66.84	2.03
Avant raffinement et avec diffusion des connaissances	62.92	1.95
Après raffinement et sans diffusion des connaissances	58.50	2.32
Après raffinement et avec diffusion des connaissances)	<u>85.74</u>	<u>1.91</u>

Comme nous pouvons le constater, les résultats obtenus dans le cas de Rayleigh sont fortement similaires au cas gaussien lorsque les classes présentent une grande complexité de discrimination. L'amélioration du taux de reconnaissance est encore plus spectaculaire suite à l'application de la diffusion croisée des connaissances possibilistes sur la base de l'ensemble des germes possibilistes capturés par le raffinement (de l'ordre de 27%). Pourtant, nous remarquons dans le cas de grande complexité de discrimination des classes (et sans l'application du processus de raffinement des distributions de possibilités) une diminution de l'ordre de 4% du taux de reconnaissance suite à l'application de la diffusion croisée des connaissances par rapport à la configuration sans diffusion des connaissances. Ce résultat, fort intéressant, signifie que la diffusion des connaissances possibilistes de qualité « médiocres ou faibles » dans le cas le plus difficile pour la classification (i.e. cas de Rayleigh avec une grande complexité de discrimination des classes), aura pour effet de dégrader les taux de reconnaissance au lieu de les améliorer. Par conséquent, l'application de la diffusion croisée des connaissances ne permet pas l'amélioration des taux de reconnaissance dans toutes les configurations des classes à discriminer. En revanche, l'application conjointe des deux processus de raffinement des distributions de possibilités et de la diffusion des connaissances possibilistes apporte une amélioration certaine du taux de reconnaissance dans tous les cas de figures.

5.3.2 Evaluation des performances au cas des classes multimodales

Dans ce paragraphe, nous proposons d'évaluer l'intérêt de l'approche ascendante proposée pour la segmentation possibiliste dans le cas où les objets recherchés dans une image sont de nature gaussienne

bimodale se trouvant dans un fond d'image de nature gaussienne. Un exemple d'images de synthèse ayant cette nature est donné dans la Figure 5.10. Dans ce cas de faible, resp. grande, complexité de discrimination de classes, l'objet d'intérêt, C_1 , est engendré par un mélange de deux gaussiennes : $G(50, 10)$ et $G(150, 10)$ Figure 5.10, resp. $G(75, 15)$ et $G(125, 15)$ Figure 5.10 ; et, la classe fond d'image, C_2 , est engendré par une loi gaussienne : $G(100, 10)$, resp. $G(100, 15)$.

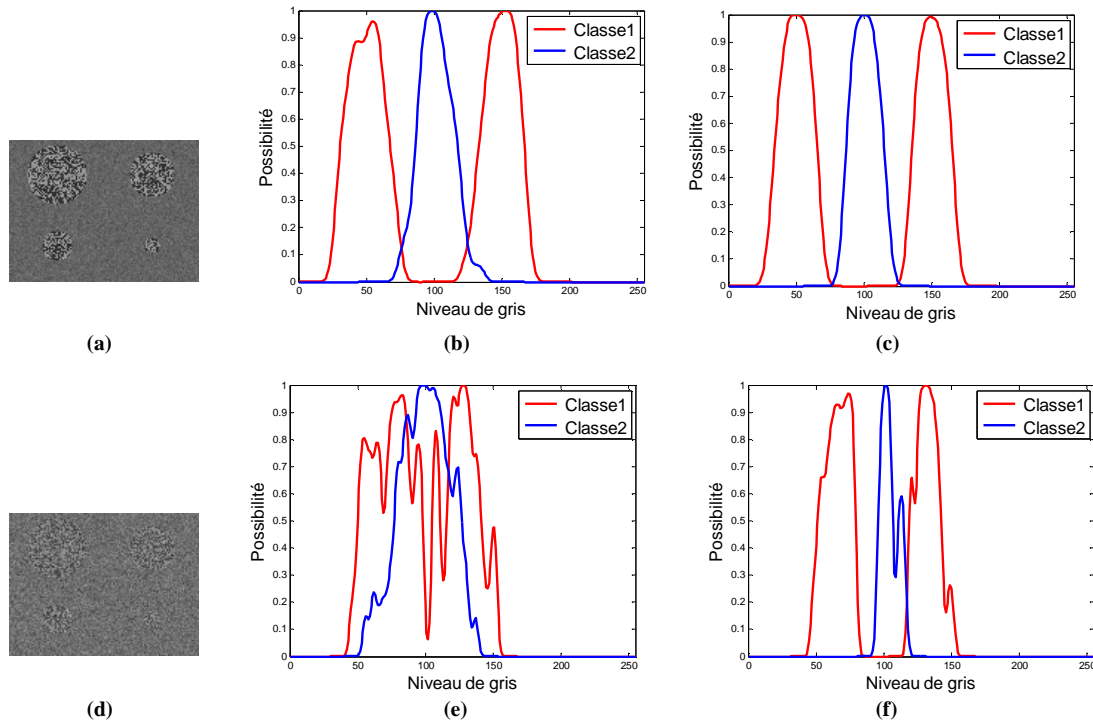


Figure 5.10 : (a), (b) et (c) Image de synthèse multimodales pour le cas de faible complexité de discrimination entre classes, distributions de possibilités initiales et raffinées respectivement des classes C_1 et C_2 et (d), (e) et (f) grande complexité de discrimination entre classes, distributions de possibilités initiales et raffinées respectivement des classes C_1 et C_2

Les distributions de possibilités initiales obtenues suite à la délimitation par l'expert de zones caractéristiques des deux classes sur les images de synthèse sont données dans la Figure 5.10, resp. Figure 5.10, dans le cas de Faible, resp. Grande, complexité de discrimination de classes.

A l'issue de l'application du processus de raffinement des distributions de possibilités initiales, nous remarquons bien que l'extraction des germes les plus fréquemment rencontrés dans les deux classes, permet de fournir l'estimation des distributions de possibilités de ces deux classes dans les deux cas considérés Figure 5.10, resp. Figure 5.10, dans le cas de Faible, resp. Grande, complexité de discrimination de classes.

La Figure 5.11 représente les cartes possibilistes et la segmentation à l'issue du processus de diffusion croisée des connaissances possibilistes et ceci pour les deux cas considérés des classes multimodales gaussiennes : faible et grande complexité de discrimination entre les classes. Les résultats obtenus montrent une homogénéité importante des régions détectées pour les deux classes C_1 et C_2 .

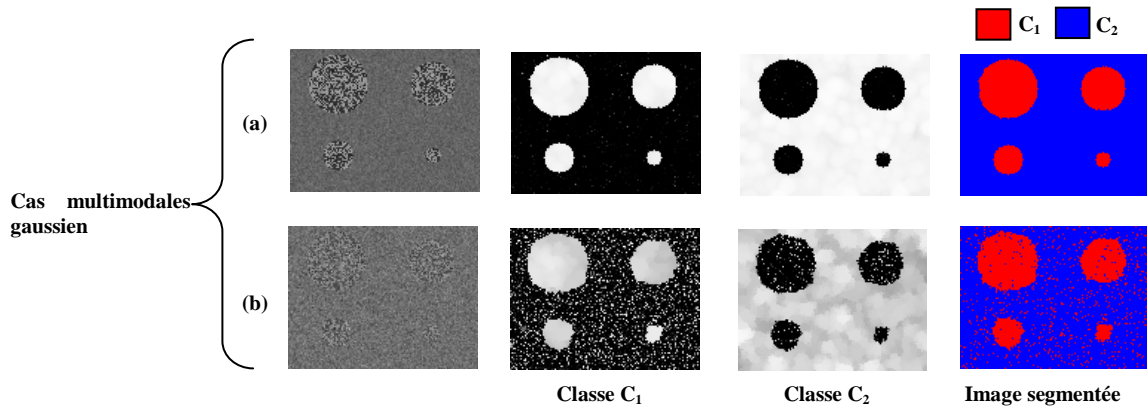


Figure 5.11 : Cartes possibilistes des classes C_1 et C_2 et les résultats de la segmentation à l'issue de la démarche de segmentation proposée (a) et (b) dans le cas multimodales gaussien avec une faible et grande complexité de discrimination entre les classes respectivement

La Figure 5.12 représente l'évolution, en fonction du nombre d'itérations, de l'erreur quadratique moyenne entre les cartes possibilistes successives obtenues au cours de la diffusion des connaissances possibilistes, dans les deux cas considérés.

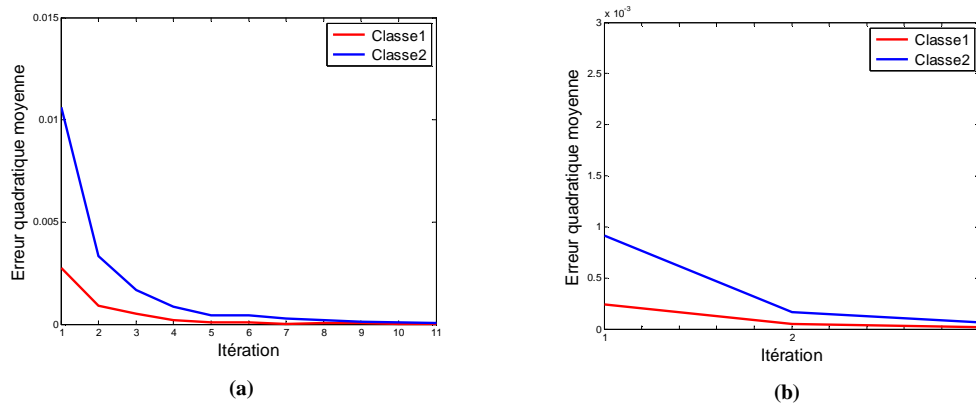


Figure 5.12 : Erreur quadratique moyenne entre deux cartes possibilistes successives dans la diffusion des connaissances possibilistes pour les classes multimodales gaussiennes (a) faible complexité de discrimination entre les classes et (b) grande complexité de discrimination entre les classes

Les résultats obtenus permettent de constater la convergence rapide du processus de diffusion croisée des connaissances possibilistes. En effet, l'erreur quadratique moyenne (et ce pour les deux classes C_1 et C_2) a atteint une valeur très proche de zéro en moins de douze itérations dans les deux cas considérés.

Concernant l'évaluation quantitative de l'approche de segmentation possibiliste proposée, une étape de classification pixelique basée sur la règle de décision de « possibilité maximale » a été appliquée sur l'image de synthèse dans les deux cas considérés des classes multimodales gaussiennes et dans les quatre configurations :

- **avant** raffinement des distributions de possibilités et **sans** diffusion des connaissances ;
- **avant** raffinement des distributions de possibilités et **avec** la diffusion croisée des connaissances possibilistes ;
- **après** raffinement des distributions de possibilités et **sans** diffusion des connaissances ;
- **après** raffinement des distributions de possibilités et **avec** la diffusion croisée des connaissances possibilistes.

Cette évaluation a été réalisée sur 100 images de synthèse (avec les mêmes paramètres pour les classes multimodales gaussiennes). Les valeurs moyennes du taux de reconnaissance ainsi que les écart-type moyens sont présentés dans le Tableau 7 (resp. Tableau 8) pour le cas de faible (resp. grande) complexité de discrimination entre les classes.

Tableau 7 : Taux de reconnaissance dans le cas des classes multimodales gaussiennes (et Faible complexité de discrimination) pour les quatre configurations considérées (avec et sans raffinement des distributions de possibilités) et (avec et sans diffusion des connaissances possibilistes)

	Taux de reconnaissance	Ecart type
Avant raffinement et sans diffusion des connaissances	98.10	0.39
Avant raffinement et avec diffusion des connaissances	99.80	0.23
Après raffinement et sans diffusion des connaissances	98.20	0.55
Après raffinement et avec diffusion des connaissances)	<u>99.95</u>	<u>0.07</u>

Dans le cas de faible complexité de discrimination entre les classes, on constate que le processus de raffinement n'a pas un grand effet sur les résultats de segmentation obtenus par la diffusion des connaissances possibilistes. En revanche, dans la situation de grande complexité de discrimination entre les classes (Tableau 8), on constate une amélioration d'environ 23% en termes de taux de reconnaissance après application conjointe du processus de raffinement des distributions de possibilités et de diffusion croisée de connaissances possibilistes par rapport au cas où seules les connaissances de l'expert sont exploitées.

Tableau 8 : Taux de reconnaissance dans le cas des classes multimodales gaussiennes (et Grande complexité de discrimination) pour les quatre configurations considérées (avec et sans raffinement des distributions de possibilités) et (avec et sans diffusion des connaissances possibilistes)

	Taux de reconnaissance	Ecart type
Avant raffinement et sans diffusion des connaissances	63.00	2.65
Avant raffinement et avec diffusion des connaissances	66.50	4.00
Après raffinement et sans diffusion des connaissances	64.32	<u>1.52</u>
Après raffinement et avec diffusion des connaissances)	<u>86.50</u>	1.84

En conclusion, dans les deux cas : monomodal et multimodal, et quelle que soit la complexité de discrimination des classes, l'application de l'approche proposée pour la segmentation possibiliste des images apporte un gain considérable en termes de taux de reconnaissance pixelique.

Les résultats obtenus par la démarche de segmentation proposée sont comparés aux résultats obtenus par l'application de la méthode de segmentation markovienne et ceci dans les différents cas considérés : cas gaussien de faible et grande complexité de discrimination entre les classes, cas de Rayleigh de faible et grande complexité de discrimination entre les classes et cas multimodal gaussien de faible et grande complexité de discrimination entre les classes (Tableau 9). Notons que la méthode de segmentation markovienne utilise l'algorithme ICM et pour les mêmes paramètres utilisés dans le chapitre précédent.

Tableau 9 : Taux de reconnaissance pixelique et écart type pour la méthode de segmentation proposée et la méthode de segmentation markovienne dans les cas suivants : cas gaussien de faible et grande complexité de discrimination entre les classes, cas de Rayleigh de faible et grande complexité de discrimination entre les classes et cas multimodal gaussien de faible et grande complexité de discrimination entre les classes

		Cas gaussien		Cas de Rayleigh		Cas multimodal gaussien	
		Faible complexité entre les classes	Grande complexité entre les classes	Faible complexité entre les classes	Grande complexité entre les classes	Faible complexité entre les classes	Grande complexité entre les classes
MRF	Taux de reconnaissance	<u>99.78</u>	86.77	<u>98.46</u>	<u>97.52</u>	93.53	78.15
	Ecart type	0.04	1.67	0.29	1.28	0.67	2.83
Approche proposée	Taux de reconnaissance	97.10	<u>89.26</u>	93.94	85.74	<u>99.95</u>	<u>86.50</u>
	Ecart type	0.09	1.40	1.27	1.91	0.07	1.84

Nous constatons que l'approche de segmentation possibiliste proposée permet d'obtenir un taux de reconnaissance proche de celui obtenu par l'approche de segmentation markovienne dans le cas gaussien de faible complexité de discrimination entre les classes. Dans les cas de Rayleigh de faible et de grandes complexités de discrimination entre les classes, l'approche de segmentation markovienne présente une supériorité par rapport à l'approche proposée. Nous remarquons aussi une supériorité de l'approche proposée par rapport à l'approche markovienne dans le cas gaussien de grande complexité de discrimination entre les classes ainsi que dans le cas multimodal gaussien de faible et grande complexités de discrimination entre les classes. Cette supériorité des résultats obtenus par l'approche proposée (par rapport à l'approche markovienne) est due de la forte dépendance de l'approche markovienne par rapport à la classification initiale réalisée. En effet, cette classification est réalisée par l'algorithme des K-moyennes qui est très sensible au bruit de type « sel et poivre ». Dans le paragraphe suivant, nous allons évaluer l'approche de segmentation possibiliste proposée sur des images réelles.

5.3.3 Evaluation des performances au cas d'images réelles

Dans ce paragraphe, l'approche proposée pour la segmentation possibiliste est appliquée sur des images réelles permettant l'observation de tumeurs de sein. Il s'agit d'images du type IRM, dans un premier temps, et des images mammographiques dans un second temps.

a) Test sur la base RIDER

Concernant les images du type IRM, la base RIDER (<http://www.cancerimagingarchive.net/>) comprenant des images IRM du sein ainsi que la vérité terrain des tumeurs (identifiée manuellement par les radiologues) est d'abord évaluée. Chaque image est composée de deux classes : tissus normaux et tissus tumoraux. Dans la Figure 5.13, deux images issues de cette base d'images ainsi que la vérité terrain (i.e. région tumorale délimitée par les radiologues) sont présentées.

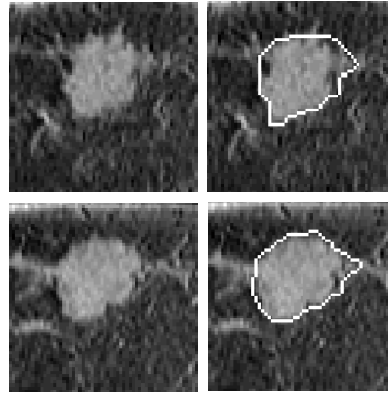


Figure 5.13 : A gauche : Régions de lésion mammaire et, à droite : la vérité terrain associée (les zones tumorales sont délimitées par des radiologues)

Deux zones d'apprentissages de taille 10×10 pixels sont positionnées par l'expert et identifiées comme étant des zones représentatives des deux classes: tissus normaux et tissus tumoraux. Ces zones d'apprentissage, les distributions de possibilités initiales (issues de ces zones d'apprentissage) ainsi que les distributions de possibilités après raffinement sont illustrées dans la Figure 5.14.

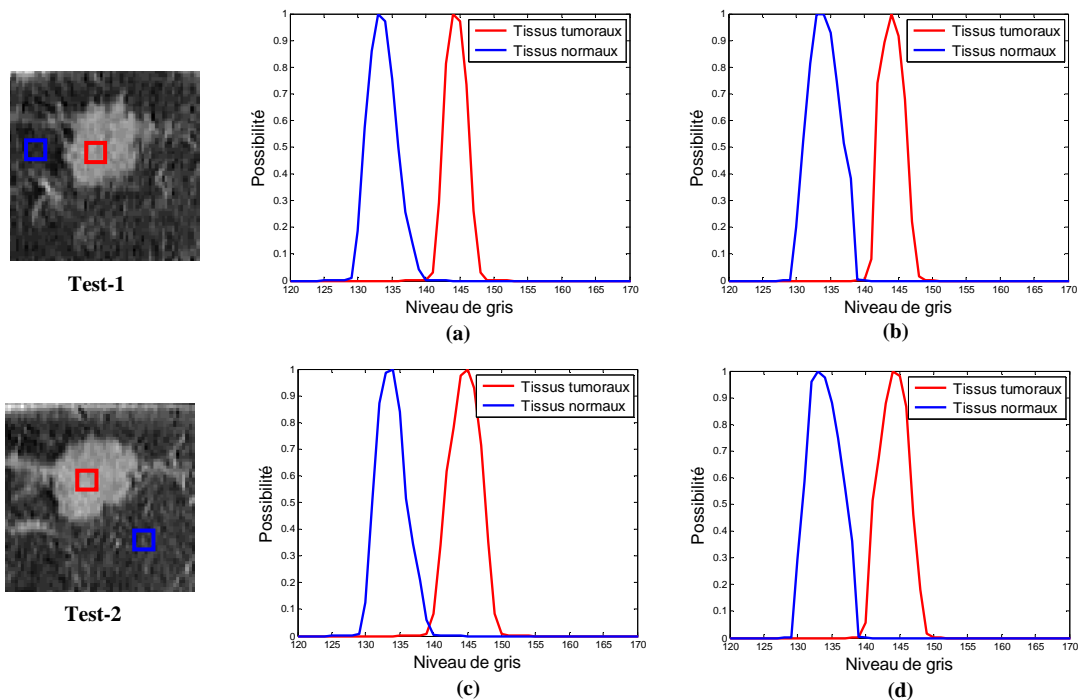


Figure 5.14 : Deux images (Test-1 et Test-2) du type IRM du sein issues de la base RIDER) avec les zones d'apprentissage délimitées par l'expert. (a) et (c) les distributions de possibilités initiales (b) et (d) et des distributions de possibilités après raffinement

Dans ce cas (considéré comme étant un cas facile de discrimination entre les classes), on constate que l'effet de raffinement est peu remarquable au niveau des distributions de possibilités initiales et raffinées. Afin d'évaluer quantitativement les résultats obtenus par la démarche de segmentation proposée, les vérités terraines des régions tumorales sont comparées aux résultats obtenus par l'application de plusieurs méthodes classiquement utilisées en segmentation des images IRM (suivies par la délimitation des régions tumorales). Trois méthodes de segmentation d'images IRM sont sélectionnées pour cette évaluation :

1. La méthode de segmentation d'image par croissance de régions [115] ;

2. La méthode de segmentation d'image semi-supervisée : Fuzzy Pattern Matching (FPM) proposée dans [110]. Cette méthode est basée sur la transformation des différents histogrammes de classe (établis à l'aide des ensembles d'apprentissage de petite taille) en des distributions de possibilités en utilisant la transformation symétrique $Pr-\pi$ de Dubois-Prade. Dans cette méthode, des échantillons « références » de chacune des classes sont utilisés pour la prise de décision afin d'attribuer un nouvel échantillon à l'une des différentes classes prédéfinies. Quand un nouvel échantillon est "accepté" comme appartenant à l'une des classes prédéfinies, cet échantillon est utilisé pour ajuster l'histogramme de la classe considérée et par conséquent, la distribution de possibilités de cette classe.
3. La méthode de segmentation markovienne utilisant l'algorithme ICM (les mêmes paramètres que ceux utilisés dans le chapitre précédent, sont appliqués).

Ces trois méthodes ont été appliquées sur les images de test données dans la Figure 5.13. Les résultats obtenus sont donnés dans la Figure 5.15, sachant que les mêmes zones d'apprentissage sont considérées pour toutes les méthodes (croissance de régions et méthode semi-supervisée FPM). Dans cette Figure 5.15, les frontières des zones tumorales identifiées (par les différentes méthodes considérées) ainsi que la réalité terrain donnée par les radiologues, sont sur imprimées. Comme on peut le constater par une simple analyse visuelle, les résultats obtenus par l'approche de segmentation proposée (en termes de délimitation des régions tumorales) sont les « plus proches » de la réalité terrain donnée par les radiologues.

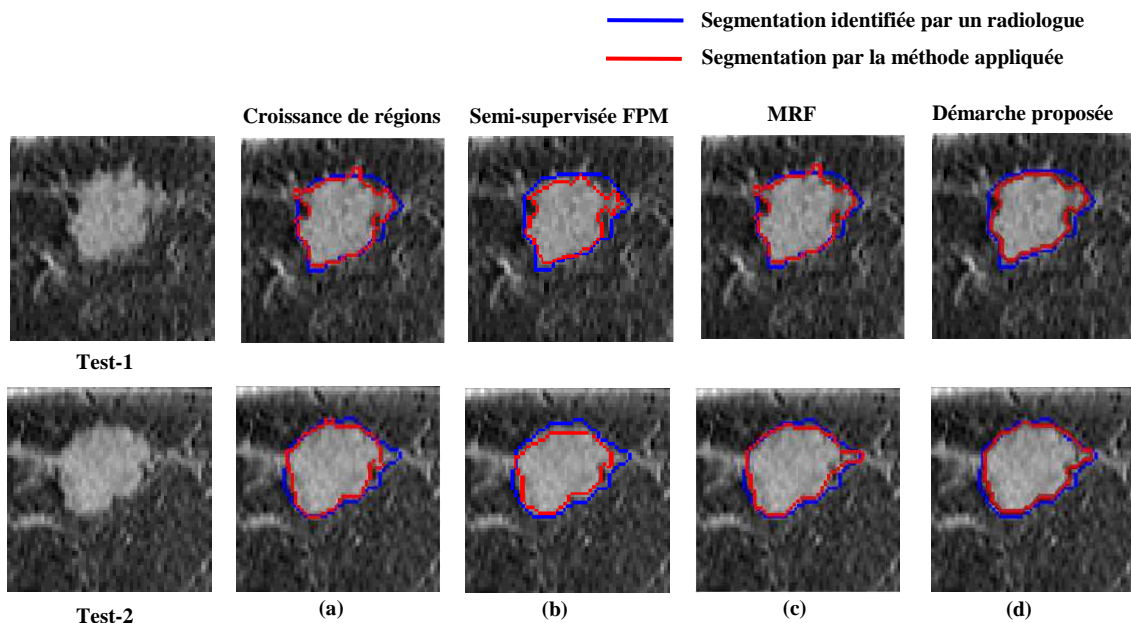


Figure 5.15 : Contour identifié de la zone tumorale et réalité terrain : (a) Méthode de segmentation par croissance de régions, (b) Méthode de segmentation semi-supervisée FPM, (c) Méthode de segmentation markovienne et, (d) Méthode de segmentation possibiliste proposée

Notons que le fait de disposer, pour la base d'images IRM RIDER, d'une réalité terrain nous permet de réaliser une évaluation quantitative du taux de reconnaissance pixelique. En effet, après avoir appliqué les quatre méthodes de segmentation considérées dans cette évaluation, nous proposons d'évaluer le taux de reconnaissance ainsi l'écart type de ce taux de reconnaissance sur un ensemble comportant 10 images issues de la base RIDER. Les résultats de cette évaluation sont présentés dans le Tableau 10.

Tableau 10 : Taux de reconnaissance pixelique et écart type pour les quatres méthodes de segmentation considérées

	Croissance de régions	Semi-supervisée FPM	MRF	Approche proposée
Taux de reconnaissance	83.95	84.74	<u>88.77</u>	87.32
Ecart type	1.76	5.35	2.53	<u>0.61</u>

Comme nous pouvons le constater, l'approche de segmentation possibiliste proposée permet d'obtenir un taux de reconnaissance proche de celui obtenu par l'approche de segmentation markovienne et supérieur aux taux de reconnaissance des deux méthodes classiquement utilisées de croissance de régions et la méthode semi-supervisée FPM. La faible valeur de l'écart type, obtenue par la démarche proposée montre également une bonne compatibilité avec le contour identifié manuellement par le radiologue.

b) Test sur la base MIAS.

Un deuxième test d'évaluation des performances de l'approche de segmentation possibiliste proposée a été réalisé sur des images mammographiques issues de la base de données MIAS (<http://www.mammoimage.org/databases/>). Cette base de données comprend des images mammographiques sans vérité terrain. Chaque image comporte deux classes de tissus : tissu normal et tissu tumoral. Deux images issues de cette base sont illustrées dans la Figure 5.16. Deux zones d'apprentissages de taille 10×10 pixels sont positionnées par l'expert comme étant des zones représentatives des deux classes considérées. Les distributions de possibilités initiales ainsi que les distributions de possibilités après raffinement sont également illustrées dans la Figure 5.16.

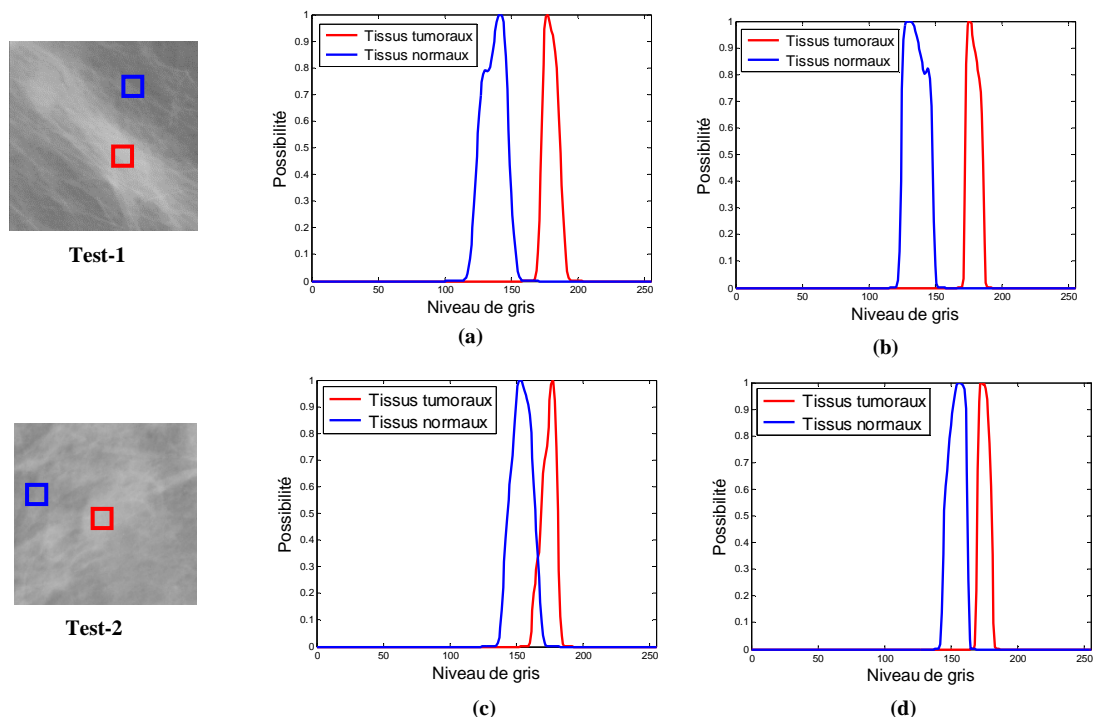


Figure 5.16 : Deux images de test mammographiques avec les zones d'apprentissage (a) et (c) les distributions de possibilités initiales (b) et (d) les distributions de possibilités après raffinement

Les quatre méthodes de segmentation d'images (appliquées pour les images du type IRM) sont évaluées pour la segmentation de ces images mammographiques. Les résultats obtenus sont donnés dans

la Figure 5.17 (les mêmes zones d'apprentissage sont prises en compte pour la méthode de croissance de régions et la méthode semi-supervisée FPM).

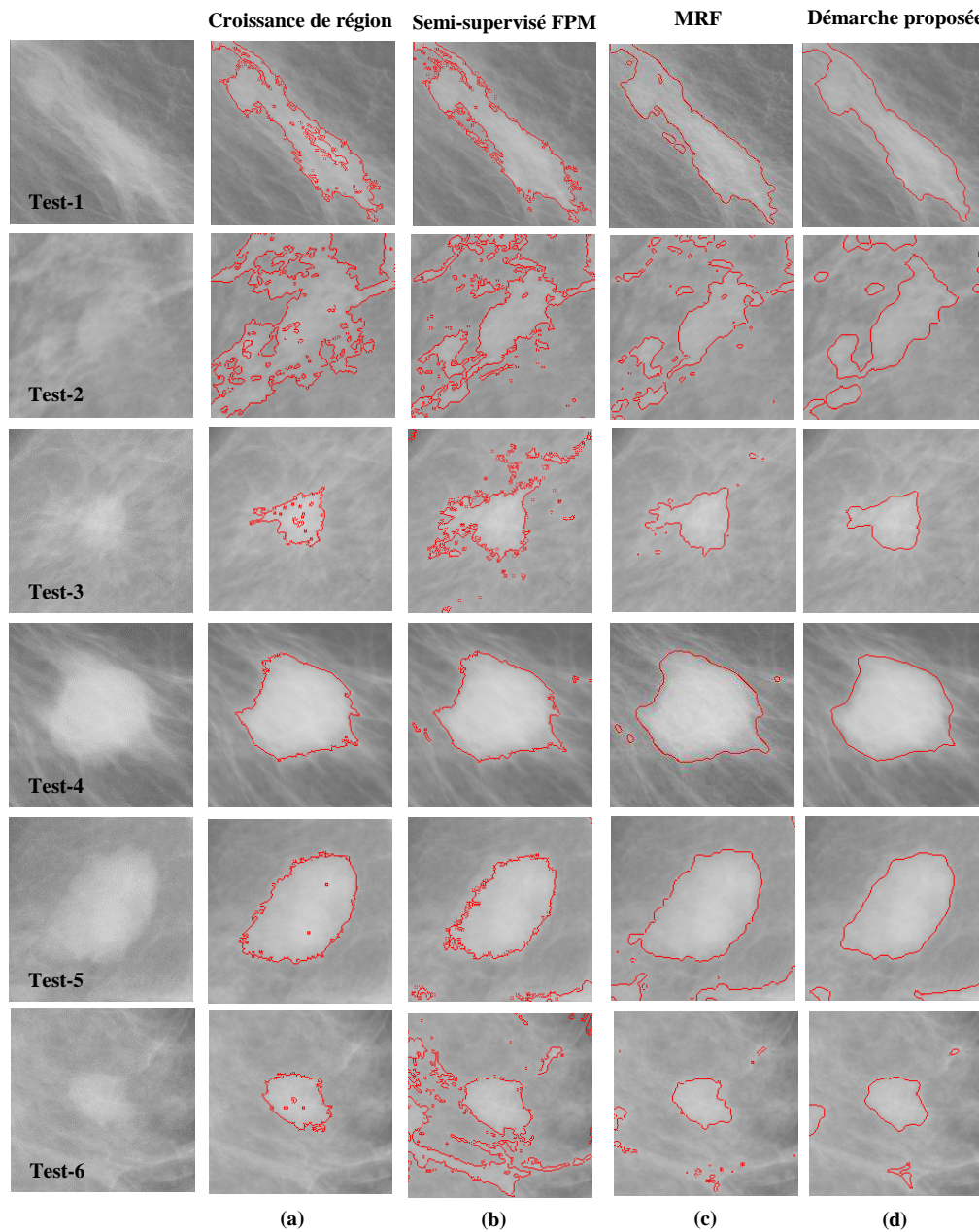


Figure 5.17 : Contour identifié de la zone tumorale: (a) Méthode de segmentation par croissance de régions, (b) Méthode de segmentation semi-supervisée FPM, (c) Méthode de segmentation markovienne et, (d) Méthode de segmentation possibiliste proposée

Une analyse visuelle des résultats obtenus montre que les performances de l'approche de segmentation possibiliste proposée sont très comparables à celles de la méthode de classification markovienne (MRF) en termes de localisation du contour. Cependant, dans la démarche proposée, le raffinement des distributions de possibilités et la diffusion possibiliste des connaissances sont des processus entièrement automatiques qui ne nécessitent aucune intervention de l'expert pour déterminer les paramètres (comme c'est le cas de la méthode de segmentation markovienne). De plus, on peut constater que la démarche proposée permet d'avoir une meilleure homogénéité des régions détectées par rapport aux régions détectées par la méthode de croissance de régions et la méthode semi-supervisée FPM.

Afin de montrer l'effet du processus de raffinement des distributions de possibilités dans l'approche de segmentation possibiliste proposée, la Figure 5.18 illustre un ensemble d'images mammographiques, issues de la base MIAS, segmentées par la méthode de diffusion croisée des connaissances possibilistes sans et avec application du processus de raffinement des distributions de possibilités.

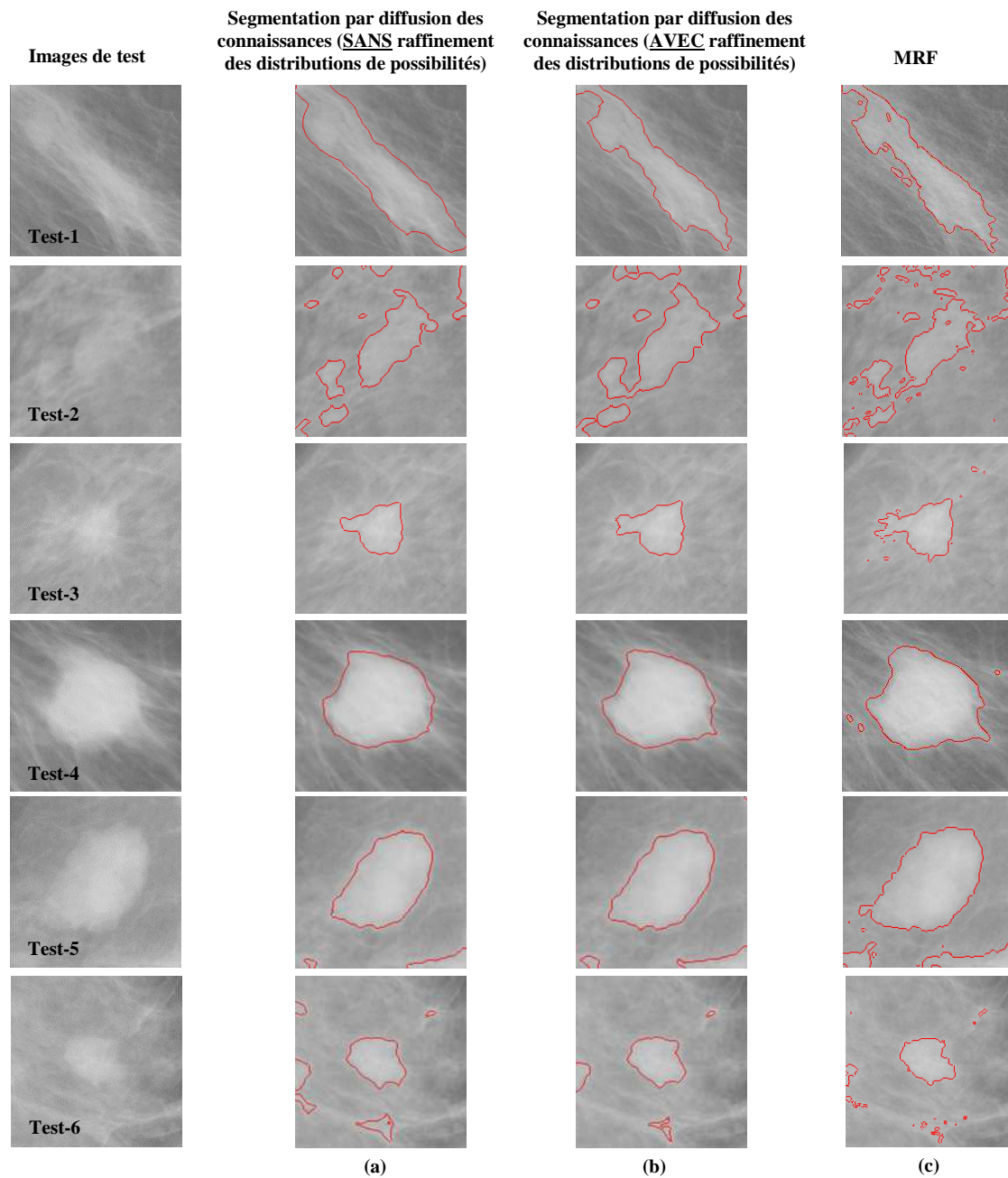


Figure 5.18 : Effet du processus de raffinement des distributions de possibilités sur les résultats de la segmentation possibiliste (a) segmentation par diffusion des connaissances et sans raffinement des distributions de possibilités, (b) segmentation par diffusion des connaissances et avec raffinement des distributions de possibilités, (c) Méthode de segmentation markovienne.

Comme nous pouvons le constater sur cette figure, l'effet du processus de raffinement des distributions de possibilités se traduit en termes d'une meilleure précision pour délimiter les zones tumorales.

Cet effet est plus marqué pour les images ayant une grande complexité de discrimination entre les classes (Test-1 et Test-2). De plus, les résultats de la segmentation obtenue sont très proches de ceux issus de la segmentation markovienne (considérée comme optimale, d'une manière générale).

Finalement, l'approche de segmentation possibiliste proposée a été appliquée sur des images SONAR représentant des fonds marins composés de deux classes : réverbération et ombre. Une particularité intéressante de ces images est que la réflectance mesurées des pixels issus de la classe réverbération (resp. ombre) sont de nature aléatoire obéissants à une loi de Rayleigh (resp. gaussienne).

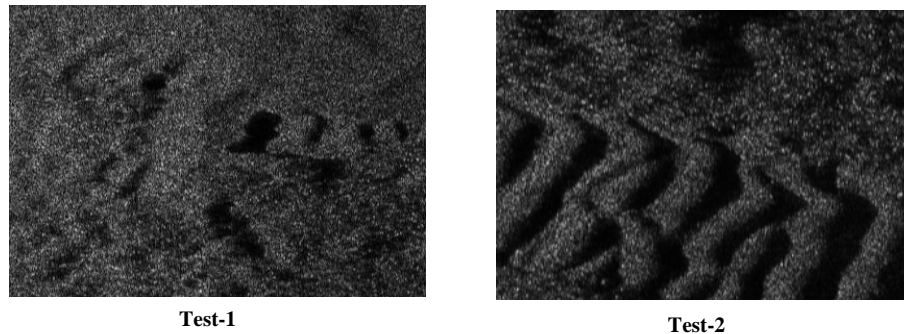


Figure 5.19 : Images SONAR utilisées pour l'évaluation de la démarche ascendante de segmentation possibiliste

Les zones d'apprentissage, délimitées par un expert, les distributions de possibilités (avant et après raffinement) ainsi que les résultats de la segmentation obtenue sont donnés dans la figure 5.20 (resp. figure 5.21) pour l'image SONAR Test-1 (resp. Test-2).

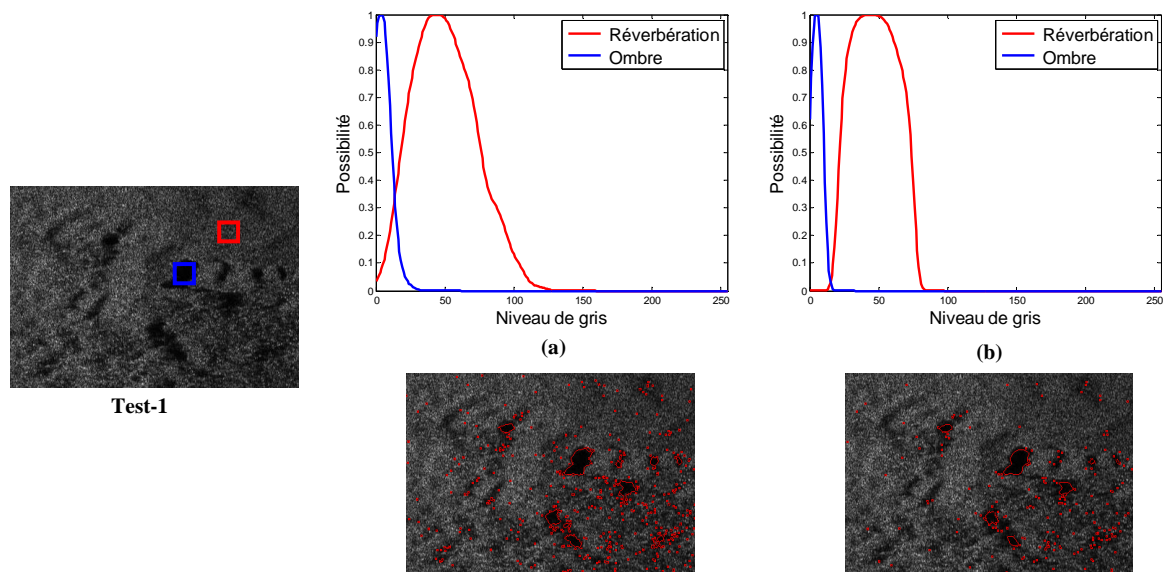


Figure 5.20 : Image SONAR Test-1 avec zones d'apprentissage,
 (a) distributions de possibilités initiales des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances ;
 (b) distributions de possibilités raffinées des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances

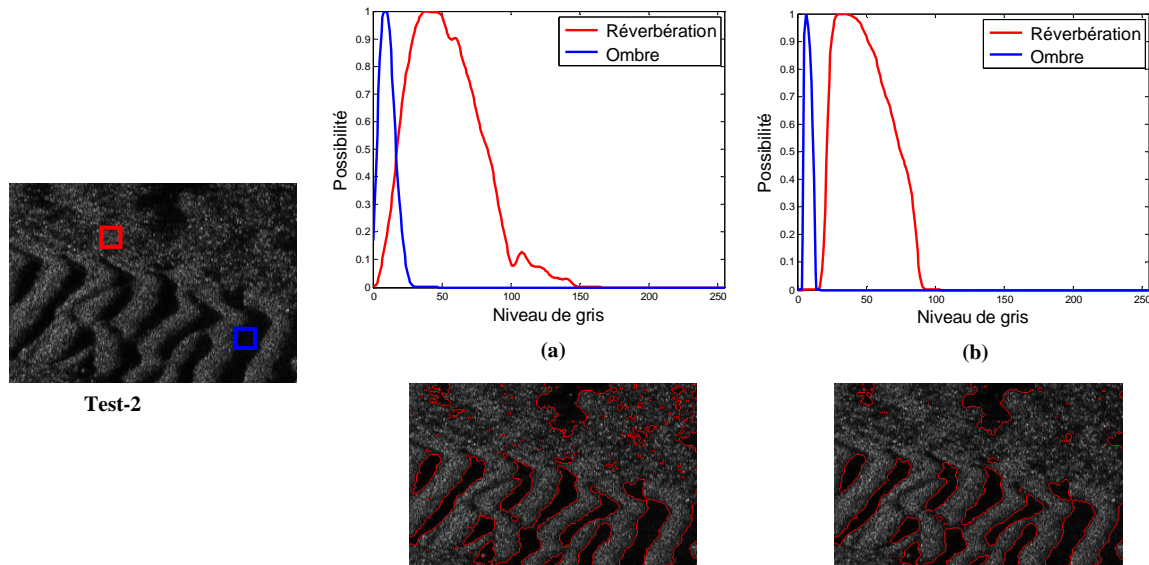


Figure 5.21 : Image SONAR Test-2 avec zones d'apprentissage,
 (a) distributions de possibilités initiales des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances ;
 (b) distributions de possibilités raffinées des deux classes (réverbération et ombre) et segmentation par diffusion des connaissances

L'analyse visuelle des résultats obtenus montre que la démarche proposée associant le processus de raffinement des distributions de possibilités avec la diffusion croisée des connaissances possibilistes permet d'avoir une meilleure homogénéité de régions détectées (par rapport aux régions détectées par le processus de diffusion sans raffinement des distributions de possibilités) et une forte diminution des zones d'ombre détectées du type fausse alarme.

Dans la suite de ce chapitre, nous allons exploiter cette démarche possibiliste de segmentation ainsi que le processus de démixage des pixels par classes (proposé dans le chapitre précédent) afin d'extraire les connaissances liées au contenu de chacune des régions résultantes du processus de segmentation.

5.4 Proposition d'une démarche « descendante » de démixage

Dans un processus d'interprétation de scène, l'intégration des connaissances descriptives de l'expert est primordiale afin d'extraire le contenu sémantique de la scène analysée. Etant donnée une représentation possibiliste de ces connaissances descriptives, l'application de l'approche descendante proposée pour la segmentation possibiliste d'une image produit des régions homogènes associées aux différentes classes thématiques identifiées par le même expert. Il s'agit donc, d'un processus de classification « possibiliste » d'une scène dans un contexte d'incertitude où on opère sous l'hypothèse du monde fermé (i.e. les M classes thématiques sont les seules classes rencontrées dans la scène analysée par l'intermédiaire de l'image) (Figure 5.22).

Notons la différence fondamentale en termes d'objectifs et de finalités entre les processus de segmentation et de classification au niveau régions. En effet, l'objectif de la segmentation consiste à identifier les régions *homogènes* dans l'image, tandis que l'objectif de la classification consiste à attribuer un label désignant l'une des classes thématiques à chaque région identifiée. Par conséquent, et sous l'hypothèse du monde fermé, le processus de classification se trouve dans l'obligation d'attribuer un unique label de classe à chaque région.

Lorsqu'il s'agit d'une région *pure* (i.e. tous les pixels de cette région représentent des cellules de résolution ne comportant que cette même classe thématique), les deux processus de segmentation et de classification sont en parfaite cohérence.

En revanche, lorsqu'il s'agit d'une région *mixte* (i.e. les pixels composants cette région sont des pixels mixtes correspondant à un mélange de classes thématiques), ou d'une région pure mais d'une classe non « prévue » *a priori* (hypothèse du monde ouvert), alors les objectifs des deux processus de segmentation et de classification présentent une forme d'incohérence au niveau de la décision qui doit être prise par le processus de classification. Deux solutions sont envisagées dans la littérature à ce problème : soit on affecte le label majoritairement présent (au niveau pixelique) à la classe considérée, soit on intègre une nouvelle classe thématique (dite : classe rejet ou classe « poubelle ») à l'ensemble des classes thématiques identifiées par l'expert.

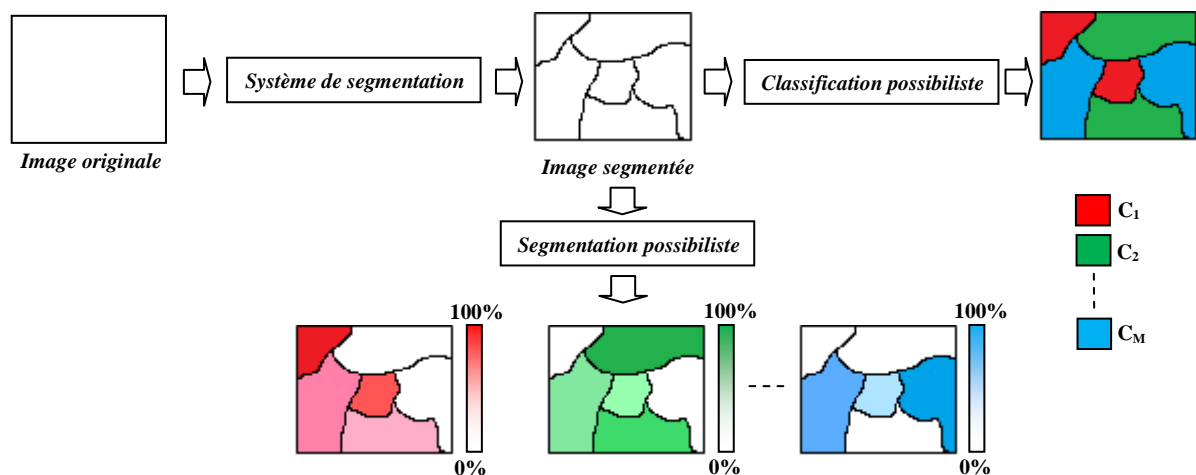


Figure 5.22 : Schéma général du système de segmentation-classification possibiliste sous l'hypothèse d'un monde fermé

Dans ce travail, nous supposons que l'on opère sous l'hypothèse d'un monde fermé comportant M classes thématiques prédéfinies par l'expert.

Afin de remédier à l'incohérence, en termes d'objectifs, entre les deux processus de segmentation et de classification, nous proposons une démarche d'interprétation descendante basée sur l'analyse de l'ensemble des pixels composants chaque région d'intérêt identifiée. La finalité de cette démarche est de produire une description plus riche en termes de connaissances qu'une simple attribution d'un label de classe à chaque région homogène extraite par l'approche de segmentation possibiliste (proposée dans le paragraphe précédente).

En effet, nous proposons l'application de la méthode de démixage pixelique basée sur la similarité possibiliste (développée dans les chapitres précédents) sur l'ensemble des pixels d'une région d'intérêt donnée afin de former un vecteur descriptif de cette région dont les composantes sont les histogrammes des taux d'abondance des classes thématiques (présentes dans la scène analysée) et calculés sur l'ensemble des pixels de la région considérée (Figure 5.23).

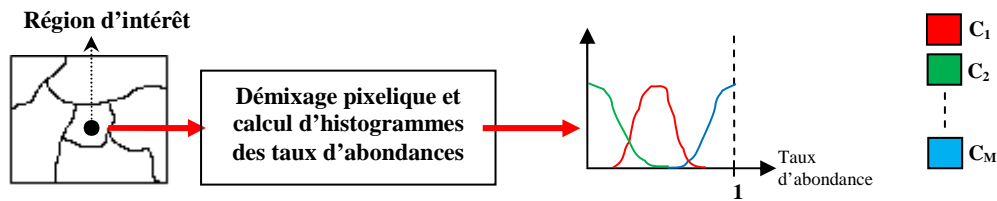


Figure 5.23 : Exemple de vecteur descriptif d'une région d'intérêt

Cette description d'une région d'intérêt est très riche, en termes d'informations, comparée à un simple processus de classification et permet de réaliser une analyse plus fine du contenu sémantique de la région considérée.

Nous allons illustrer cette démarche en s'appuyant sur l'image de synthèse donnée dans la Figure 5.24 comportant deux classes thématiques C_1 engendrée par une loi gaussienne $G(m_1, \sigma_1)$, et C_2 engendrée par une loi gaussienne $G(m_2, \sigma_2)$. Quatre objets de forme circulaire constituant les régions d'intérêt sont présentes dans cette image :

- une région R_1 (en haut à gauche) représentant un objet pur de la classe C_1 ;
- une région R_2 (en bas à gauche) représentant un objet composé de pixels mixtes qui sont un mélange linéaire : $A_1 = 0.8$ de la classe C_1 et $A_2 = 0.2$ de la classe C_2 ;
- deux régions R_3 et R_4 (en haut à droite et en bas à droite) représentant deux objets composés de pixels mixtes qui sont un mélange linéaire : $A_1 = 0.6$ de la classe C_1 et $A_2 = 0.4$ de la classe C_2 .

De plus, deux cas sont évalués : un cas où les deux classes C_1 et C_2 sont de faible complexité à être discriminées ($C_1 \equiv G(160, 10)$ et $C_2 \equiv G(100, 15)$), et un cas où les deux classes C_1 et C_2 sont de grande complexité à être discriminées ($C_1 \equiv G(120, 15)$ et $C_2 \equiv G(100, 20)$).

Rappelons que les connaissances d'expert sont données par la délimitation des zones d'apprentissage de taille 10×10 pixels.

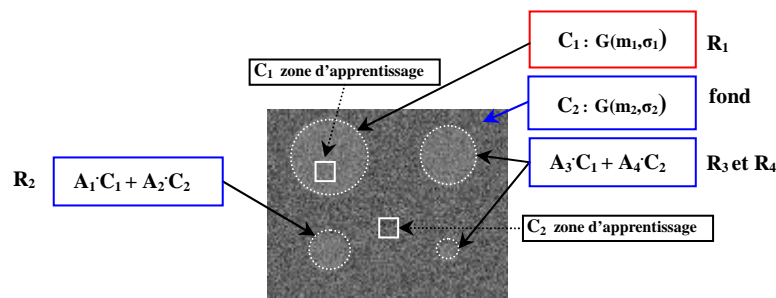


Figure 5.24 : Image de synthèse-type avec les zones d'apprentissage

Les résultats de la segmentation par l'approche possibiliste proposée, pour les deux images de synthèse, sont donnés dans la Figure 5.25. Par simple analyse visuelle de ces résultats, nous constatons que la démarche possibiliste de segmentation proposée permet de détecter les régions d'intérêts avec, dans le cas de grande complexité de discrimination, quelques petites régions de type fausse alarme.

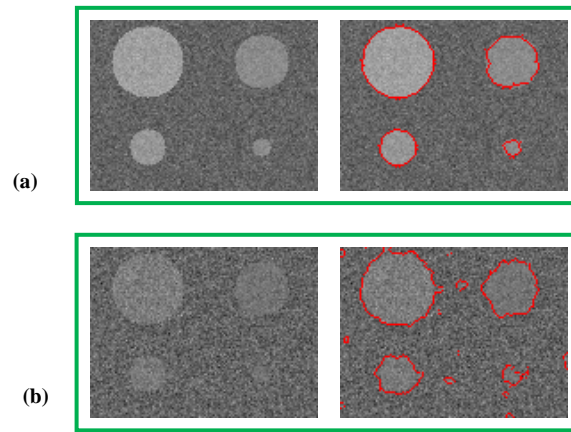
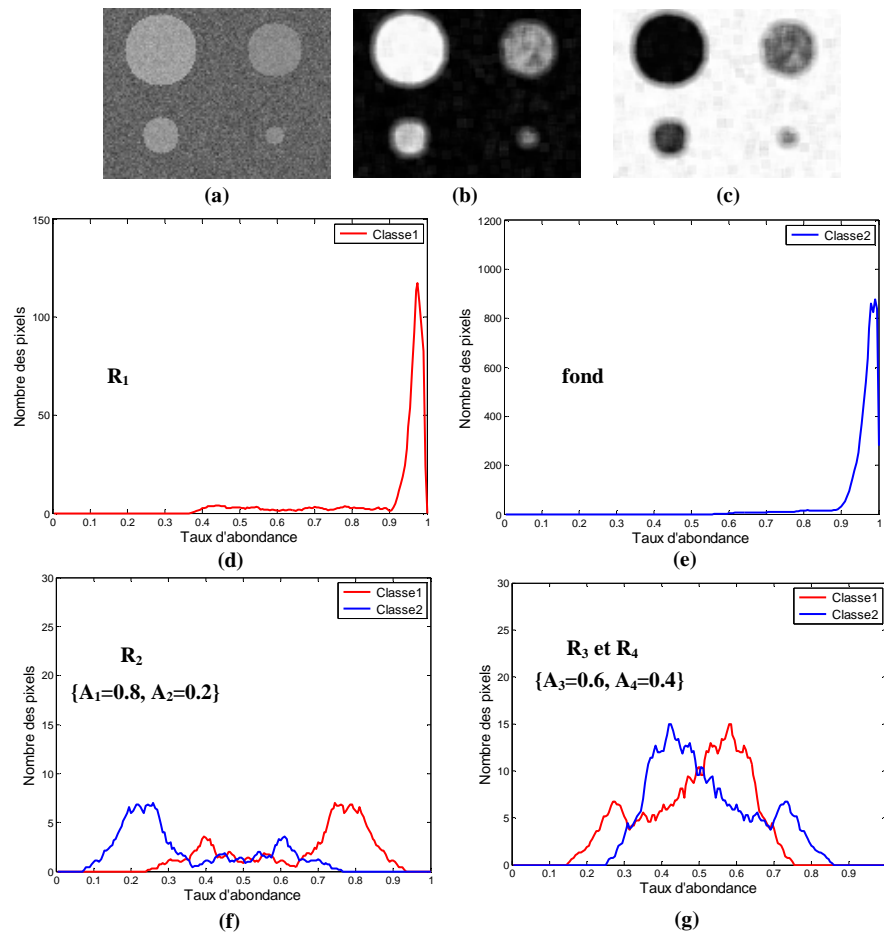


Figure 5.25 : Images de synthèse segmentées par l'approche possibiliste proposée : (a) cas de faible complexité de discrimination, et (b) cas de grande complexité de discrimination

Le processus de démixage pixelique, basé sur la similarité possibiliste, est ensuite appliqué sur l'image de synthèse dans les deux cas considérés (des fenêtres de taille 5×5 autour de chaque pixel de l'image sont considérées). Les cartes des taux d'abondance pour les deux classes C_1 et C_2 ainsi que les vecteurs descriptifs des régions d'intérêts détectées (en phase de segmentation possibiliste) sont illustrés dans la Figure 5.26 et la Figure 5.27. L'inspection visuelle des cartes d'abondances (dans les deux cas considérés) montre une excellente correspondance entre ces cartes et la réalité-terrain de ces images de synthèse.

Dans le cas de faible complexité de discrimination, nous remarquons la parfaite adéquation des histogrammes des taux d'abondance des classes pures (région R_1 correspondant à la classe C_1 , le fond de l'image correspondant à la classe C_2). De plus, les histogrammes des taux d'abondance des classes C_1 et C_2 dans la région d'intérêt R_2 (resp. régions R_3 et R_4) présentent des formes monomodales centrées sur les vrais coefficients de mélange de classes $A_1 = 0.8$ et $A_2 = 0.2$ (resp. $A_1 = 0.6$ et $A_2 = 0.4$).



**Figure 5.26 : (a) Image de synthèse (cas de faible complexité de discrimination),
 (b) et (c) les cartes des taux d'abondance pour les deux classes C_1 et C_2 respectivement,
 (d) (e) les vecteurs descriptifs des taux d'abondance des classes pures (région R_1 correspondant à la classe C_1 , le fond de l'image correspondant à la classe C_2)
 (f) et (g) les vecteurs descriptifs des taux d'abondance des classes C_1 et C_2 dans la région d'intérêt R_2 (resp. régions R_3 et R_4)**

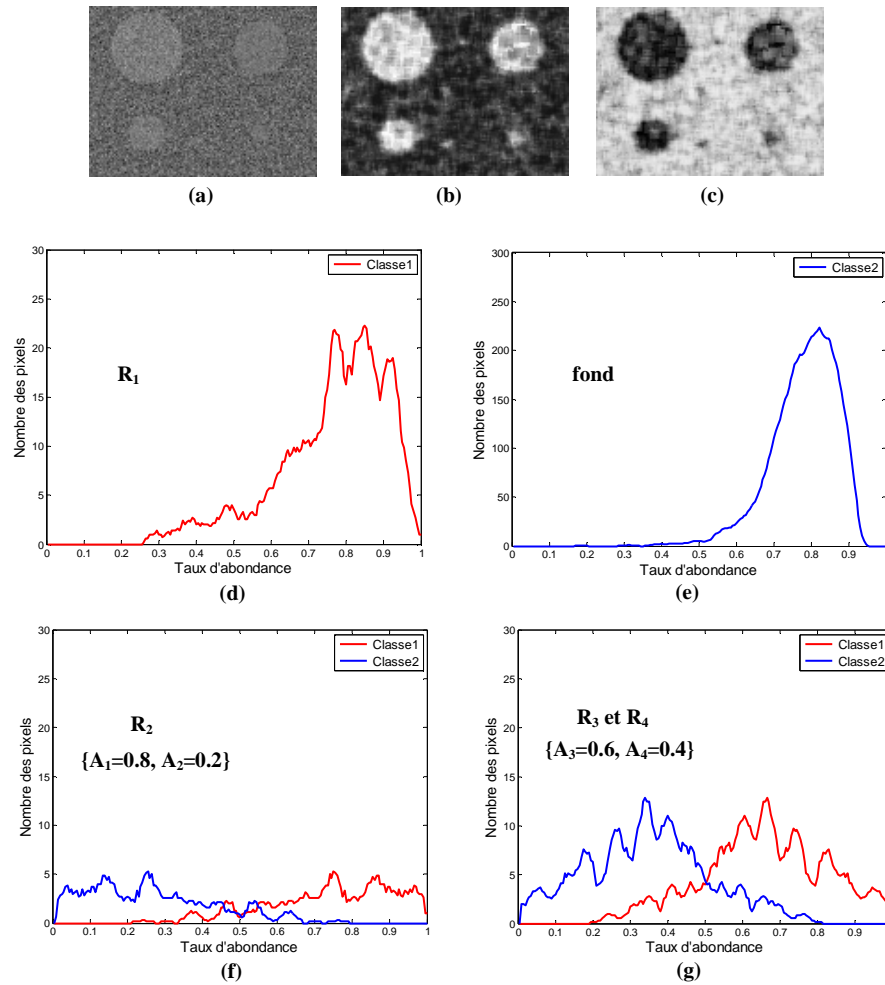


Figure 5.27 : (a) Image de synthèse (cas de grande complexité de discrimination), (b) et (c) les cartes des taux d'abondance pour les deux classes C_1 et C_2 respectivement, (d) et (e) les vecteurs descriptifs des taux d'abondance des classes pures (région R_1 correspondant à la classe C_1 , le fond de l'image correspondant à la classe C_2) (f) et (g) les vecteurs descriptifs des taux d'abondance des classes C_1 et C_2 dans la région d'intérêt R_2 (resp. régions R_3 et R_4)

Malgré la grande complexité de discrimination des deux classes C_1 et C_2 , nous remarquons que les histogrammes des taux d'abondance des classes pures (région R_1 correspondant à la classe C_1 , le fond de l'image correspondant à la classe C_2) suggèrent fortement la pureté (i.e. non ou très faible mélange) des deux classes C_1 et C_2 . De plus, et malgré cette complexité de discrimination, les histogrammes des taux d'abondance des classes C_1 et C_2 dans les régions d'intérêt R_2 , R_3 et R_4 permettent d'apprécier les vrais coefficients de mélange de classes (les moyennes de ces histogrammes correspondent aux vraies valeurs de ces coefficients : $A_1=0.8$ et $A_2=0.2$ pour la région R_2 et $A_1=0.6$ et $A_2=0.4$ pour les régions R_3 et R_4).

5.4.1 Application de l'approche descendante aux images réelles

Dans ce paragraphe, nous proposons d'appliquer la démarche possibiliste proposée pour l'interprétation de régions d'intérêt issues de la segmentation possibiliste dans le cas d'images mammographiques, extraite de la base d'images MIAS fournie par la société savante anglaise « *Mammographic Image Analysis Society* ». Notons que dans ces images, les régions d'intérêt correspondent à des zones de surdensité dans la glande mammaire. En effet, le médecin focalise son attention sur ces zones qui comportent des tissus tumoraux potentiels et il souhaite être conforté dans sa décision.

L'image mammographique donnée dans la Figure 5.28-(a) illustre un exemple d'une image comportant une zone tumorale faiblement texturée et de forte contraste. Cette image est considérée par les médecins, comme étant d'une faible complexité de discrimination de la zone tumorale par rapport au tissu sein. Le résultat de l'application de l'approche ascendante proposée pour la segmentation possibiliste est illustré dans la Figure 5.28-(b) (seuls les contours des zones tumorales sont portés sur cette figure). L'application de l'approche descendante (en utilisant une fenêtre démixage de taille 5×5) pour l'interprétation de la zone tumorale centrale aboutit aux cartes de taux d'abondance pour la classe Tissu tumoral et Tissu sein données dans La Figure 5.28-(c) et la Figure 5.28-(d) respectivement. L'histogramme de taux d'abondance de la classe Tissu tumoral est donné dans la Figure 5.28-(e). Comme nous pouvons le remarquer, le taux d'abondance pour cette classe est fortement concentré autour de la valeur unitaire maximale. L'histogramme de taux d'abondance pour la classe Tissu normal est aussi donné dans la zone de tissu sein (Figure 5.28-(f)). Finalement, l'histogramme de taux d'abondance pour la classe Tissu tumoral est donné pour une zone à cheval sur les deux types de tissu, Figure 5.28-(g). Nous remarquons la forte baisse du taux d'abondance à la classe de tissu tumoral due à la présence du tissu normal dans cette fenêtre.

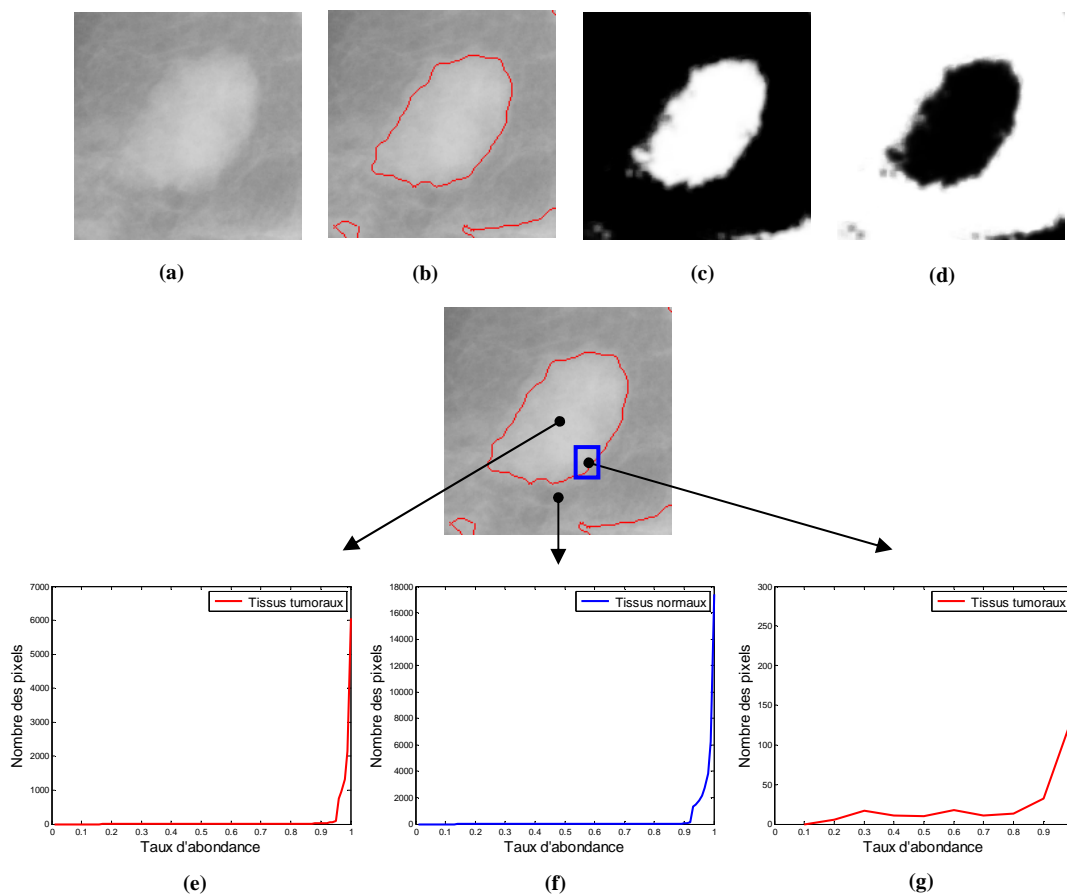


Figure 5.28 : Approche descendante appliquée pour l'interprétation d'images mammographiques (a) Image originale (cas de faible complexité de discrimination) ; (b) Image segmentée par la démarche possibiliste proposée ; (c) & (d) les cartes des taux d'abondance des deux classes « Tissu tumoral » et « Tissu sein » respectivement ; (e) Histogramme d'abondance de la classe Tissu tumoral de la zone suspecte ; (f) Histogrammes de taux d'abondance des deux classes : Tissu normal & Tissu tumorale (dans la zone de tissu sein) ; (g) Histogramme de taux d'abondance de la classe Tissu tumoral dans une fenêtre comportant les deux types de tissu

L'image mammographique donnée dans la Figure 5.29 illustre un autre exemple d'une image considérée par le médecin comme étant de grande complexité. Cette image comporte une zone tumorale

fortement texturée, de faible contraste et de multiple régions pour lesquelles le médecin porte un intérêt particulier afin d'étudier l'étendue de la zone tumorale. Les résultats obtenus aboutissent aux mêmes conclusions que celles de l'image considérées dans la Figure 5.28. L'analyse du résultat de l'application de l'approche descendante à une région détectée par l'approche ascendante, Figure 5.29-(f), montre clairement qu'il s'agit d'une région comportant une forte concentration de tissu tumoral montrant ainsi la présence d'autres foyers tumoraux dans le voisinage de la zone tumorale centrale.

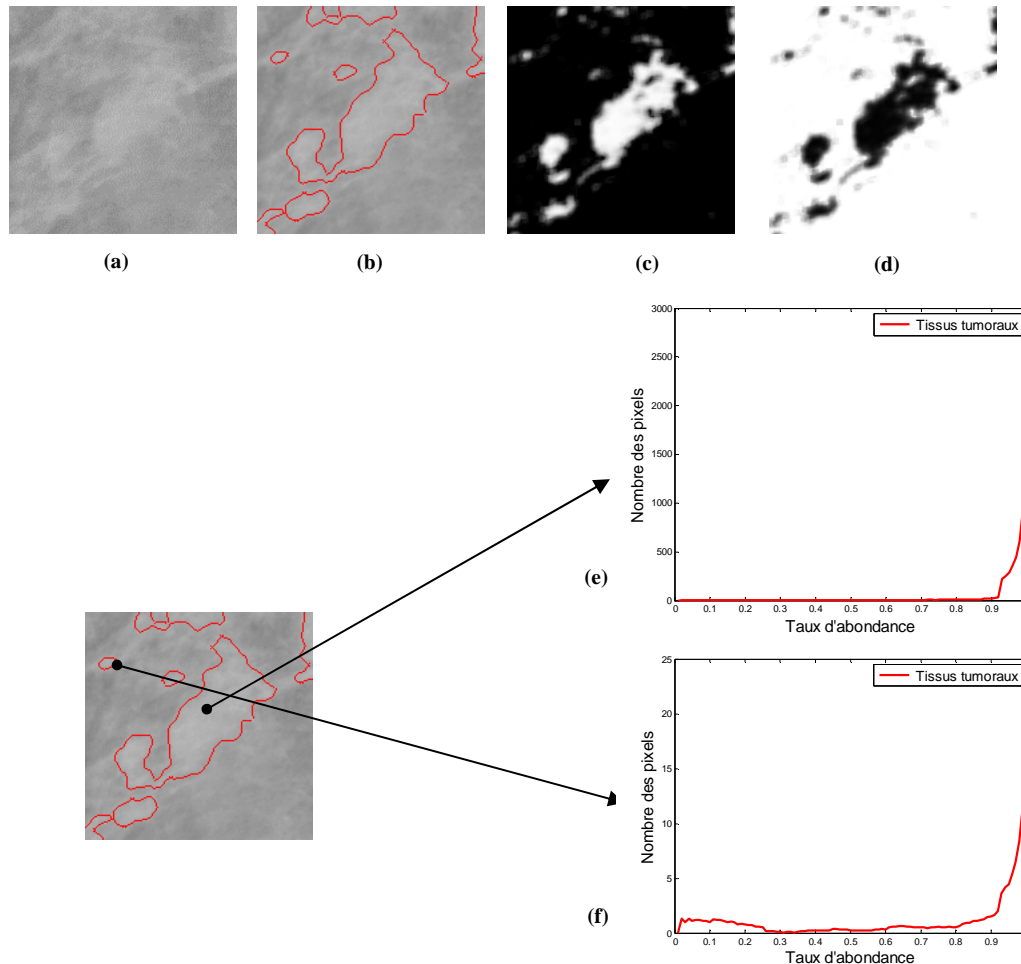


Figure 5.29 : Approche descendante appliquée pour l'interprétation d'images mammographiques (a) Image originale (cas de grande complexité de discrimination) ; (b) Image segmentée par la démarche possibiliste proposée ; (c) & (d) les cartes des taux d'abondance des deux classes « Tissu tumoral » et « Tissu sein » respectivement ; (e) Histogramme de taux d'abondance de la classe Tissu tumoral de la zone suspecte ; (f) Histogramme de taux d'abondance de la classe Tissu tumoral dans une masse suspecte

Par conséquent, il est clair que les vecteurs descriptifs, proposés par l'approche descendante et extraits à partir des régions détectées, permettent aux médecins d'avoir une idée plus précise concernant le contenu informationnel de ces régions.

5.5 Conclusion

Dans ce chapitre, nous avons proposé une approche globale permettant de contribuer aux objectifs visés par l'interprétation de scène. Les approches proposées de raffinement des distributions de possibilités, de segmentation par diffusion croisée des connaissances possibilistes ainsi que de démixage

pixelique possibilistes sont cadencées séquençement dans l'objectif de former l'approche globale proposée. L'approche d'interprétation de scène proposée comporte deux phases : une phase ascendante et une phase descendante.

La phase ascendante consiste à considérer les connaissances injectées par l'expert comme un point de départ. Ces connaissances sont d'abord transformées en des distributions de possibilités initiales. Ensuite, un processus de raffinement de ces distributions est appliqué. L'objectif du raffinement consiste à extraire le maximum d'échantillons caractéristiques, les plus fréquents, afin de réaliser à partir de ces échantillons, une nouvelle estimation des distributions de possibilités des différentes classes présentes dans la scène analysée. Finalement, la phase ascendante opère une segmentation possibiliste par la diffusion des connaissances possibilistes qui résultent du processus de raffinement des distributions de possibilités. Ces connaissances sont présentées sous la forme de cartes possibilistes d'appartenance aux différentes classes. A l'issue de la phase ascendante, la scène analysée est représentée sous la forme d'une carte de régions homogènes.

La phase descendante de l'approche proposée consiste à réaliser l'interprétation des différentes régions homogènes identifiées dans la phase ascendante par le biais de l'approche d'analyse et de démixage pixelique possibiliste proposée dans les chapitres précédents. L'objectif de cette analyse consiste à représenter chaque région par un vecteur de descripteurs comportant les histogrammes des taux d'abondance aux différentes classes de la scène analysée. Un vecteur de descripteurs apporte une source de connaissance très riche permettant d'interpréter d'une façon fine le contenu sémantique de chaque région en termes de composition pixelique-thématique de l'ensemble des pixels composant cette région. L'objectif visé par cette phase descendante consiste soit à confirmer la présence d'une unique classe thématique dans chaque région, soit à analyser son contenu sémantique afin qu'un expert ou un système automatisé de prise de décision puisse prendre une décision finale soit en attribuant la région à une seule classe, soit en explicitant les différentes classes formant cette région ainsi que leur taux d'abondance.

Nous avons évalué l'évaluation des performances de l'approche ascendante de segmentation possibiliste dans le cas de classes monomodales et multimodales. Les résultats obtenus sur la base d'images de synthèse ont montré l'effet de regroupement de l'approche de raffinement des distributions de possibilités et la segmentation par diffusion croisée des connaissances possibilistes et ceci en termes d'augmentation du taux de reconnaissance surtout dans un contexte où les classes thématiques sont difficiles à discriminer.

Nous avons aussi évalué la performance de la démarche possibiliste de segmentation d'image en termes de convergence. Les résultats obtenus ont montré une convergence rapide du processus de diffusion des connaissances possibilistes (moins de quatorze itérations) et ceci dans tous les cas considérés.

Les résultats de l'évaluation montrent l'efficacité de l'approche proposée en termes d'interprétation (identification du contenu thématique des régions d'intérêt), sur la base d'images de synthèse (dont on maîtrise parfaitement bien leur contenu) ainsi que sur des images médicales liées à l'analyse de tumeurs de sein (essentiellement des images mammographiques).

Chapitre 6

VERS UN SYSTEME GLOBAL D'INTERPRETATION DE SCENE

6.1 Introduction

Dans le chapitre précédent, nous avons proposé une approche de segmentation/interprétation d'image basée sur une démarche possibiliste comportant deux phases: ascendante et descendante.

La phase ascendante est composée de deux étapes consécutives : le *raffinement* et la *diffusion des connaissances possibilistes*. Les connaissances de l'expert (i.e. connaissances « haut niveau »), liées aux différentes classes de la scène analysée, constituent le point de départ de cette phase. Un traitement collaboratif entre ces connaissances haut niveau et les données issues du capteur bas niveau permet d'accumuler les évidences sur les régions (ou objets) et aboutit à une image segmentée-classifiée de la scène analysée.

La phase descendante consiste à réaliser l'interprétation des différentes régions homogènes identifiées dans la phase ascendante. Dans cette phase, l'approche de démixage pixelique possibiliste proposée dans le quatrième chapitre est appliquée au niveau des régions permettant d'analyser l'ensemble des pixels composant chaque région d'intérêt identifiée. La finalité de cette analyse est de caractériser chaque région d'intérêt par un vecteur de descripteurs comportant les histogrammes des taux d'abondance aux différentes classes de la scène analysée. Ainsi, en explicitant les différentes classes formant chaque région, une décision peut être prise concernant l'homogénéité de la région d'intérêt (i.e. une région homogène *pure* ou *mixte*).

L'approche de segmentation/interprétation d'image proposée a été jusque-là basée essentiellement sur une seule primitive (i.e. raffinement des distributions de possibilités décrivant chaque classe) extraite

d'une seule image (i.e. une seule source de données). De plus, notre attention a été focalisée sur l'interprétation de la région d'intérêt donnée.

En effet, le système d'interprétation de scène vise à comprendre la totalité de la scène observée (identifier et analyser les différentes régions ou objets composant la scène et étudier leurs relations spatiales). Dans la conception de tels systèmes, on constate une tendance croissante d'exploiter plusieurs primitives afin de réduire la confusion interclasse. Cette tendance est motivée par l'accessibilité de plusieurs sources de données (par exemple, images visible, infrarouge, radar, SIG, ...). Les primitives utilisées peuvent être calculées à partir d'une seule image (niveau de gris, écart type, entropie, etc.) ou en provenance de plusieurs images issues de différents capteurs (après une phase de recalage géométrique de ces images).

Dans ce chapitre, nous proposons d'étendre l'approche de segmentation/interprétation d'image proposée dans le chapitre précédent et ceci afin de doter cette approche d'opérations de fusion permettant d'exploiter plusieurs primitives dans la phase ascendante, mais aussi d'un système d'analyse de la scène dans la phase descendante (Figure 1.3).

Dans la phase ascendante, deux stratégies de traitement seront considérées. La première consiste à réaliser le traitement capteur par capteur (ou image par image), d'une façon indépendante et ensuite à réaliser la fusion des résultats. L'autre stratégie consiste à réaliser le traitement conjointement en prenant en compte tous les capteurs ou toutes les images. Ces deux stratégies aboutissent chacune à une image segmentée qui constitue l'entrée du système d'analyse de la scène.

La première phase du système d'analyse de scène consiste en la construction d'un rapport descriptif pour chacune des régions identifiées dans la phase ascendante. Le rapport descriptif d'une région comporte des informations liées à cette région (le vecteur de descripteurs, la forme et la taille, etc.) ainsi que des informations liées aux relations spatiales entre la région d'intérêt et les autres régions détectées (régions adjacentes par exemple). Dans le processus de construction des rapports descriptifs et afin de former le vecteur des descripteurs de chaque région, l'approche de démixage pixelique possibiliste proposée sera appliquée sur chacune des régions identifiées. En exploitant ces rapports descriptifs, une phase d'analyse de la scène par région sera ensuite menée dans l'objectif de localiser les régions homogènes « pures » et de fusionner des régions homogènes « mixtes » (par « *Region Merging* ») ayant des caractéristiques similaires selon des critères prédéfinis et finalement donner une interprétation globale de la scène analysée.

Une évaluation globale de l'approche proposée sera détaillée et appliquée à une image comportant plus de deux classes.

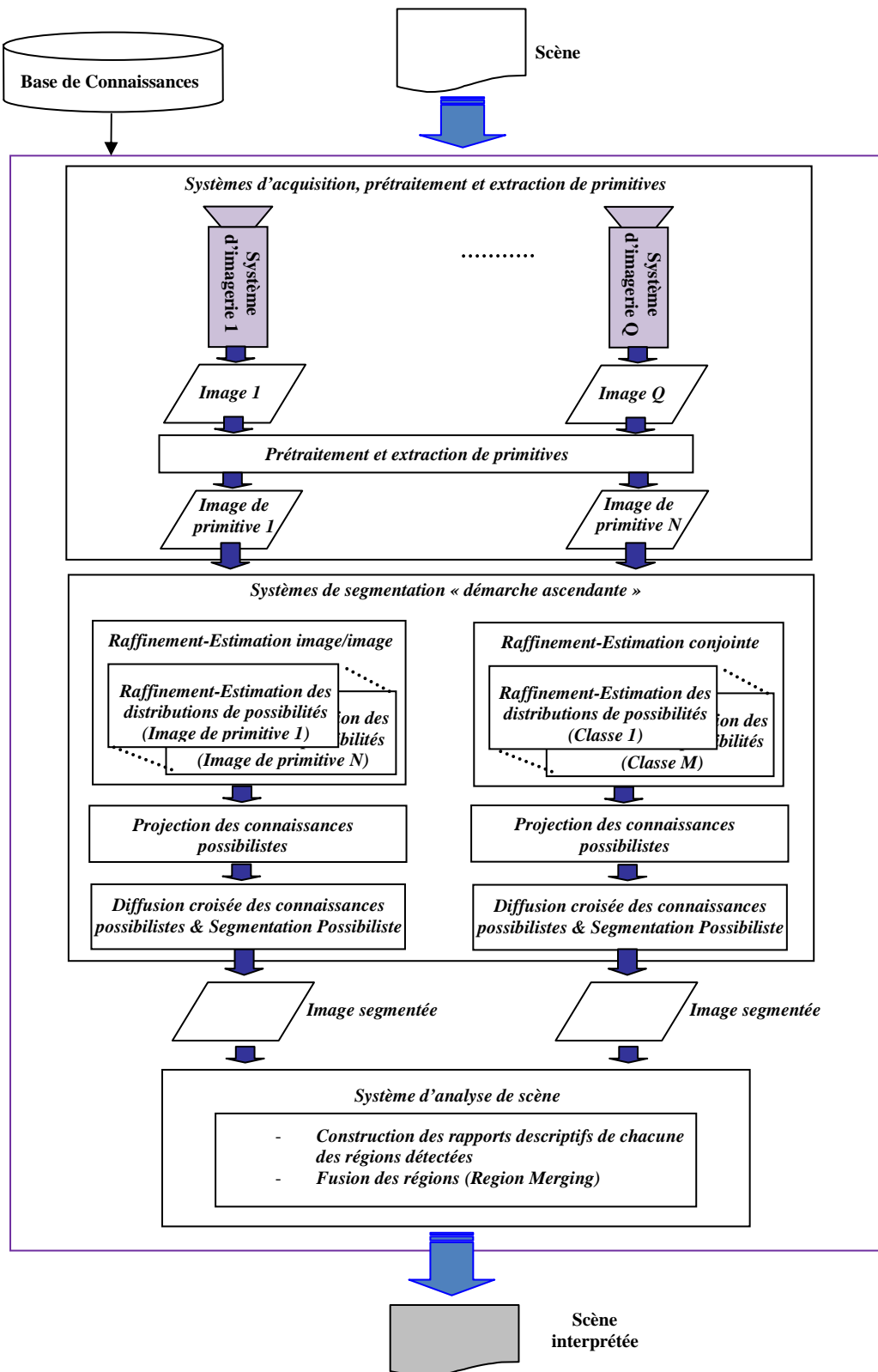


Figure 6.1 : Schéma général du système proposé d'interprétation de scène

6.2 Fusion inter-primitives

La sélection des primitives constitue une étape très importante dans le processus d'interprétation de scène. À ce stade, il faut rappeler que l'utilisation des primitives corrélées (qui ne représentent pas la variation entre les classes) peut dégrader plutôt que d'enrichir les résultats à l'issue du processus d'interprétation [63]. Notons qu'une pondération peut être attribuée à chacune de ces primitives (ou aux sources de données considérées) afin de prendre en compte la fiabilité de ces primitives (ou les sources). Cette pondération doit être aussi choisie d'une façon judicieuse afin d'éviter la dégradation dans le processus d'interprétation.

L'objectif dans ce paragraphe est d'étudier l'intérêt de fusionner plusieurs images de primitives dans l'approche de segmentation/interprétation d'image proposée et ceci en exploitant les opérateurs de fusion proposés dans le cadre de la représentation possibiliste. Notons que les images de primitives peuvent être extraites d'une seule image de la scène (observée à partir d'un seul capteur) ou à partir de plusieurs images issues de plusieurs capteurs (telle que les images multi-spectrales, les images hyper-spectrales, ...). D'une manière générale, la sélection des primitives calculées à partir de plusieurs images recalées (issues de différents capteurs) permet d'obtenir une complémentarité spectrale et spatiale plus importante que celle obtenue à partir d'une seule image [63] [71].

Dans un premier lieu, nous considérons la scène analysée (et observée via Q images recalées I_1, I_2, \dots, I_Q ayant la même taille $H \times W$) comme étant formée de M classes thématiques $\psi = \{C_1, C_2, \dots, C_M\}$ prédéfinies par l'expert qui délimite aussi un ensemble comportant des zones d'apprentissage initiales pour ces différentes classes thématiques dans les images $I_q, q=1, \dots, Q$ données de la scène analysée. Supposons que N images de primitives ont été extraites à partir des Q images originales de la scène. Notons que les Q images originales comportant les données brutes en sortie des capteurs peuvent être un cas particulier d'images de primitives ($Q = N$).

Dans la phase ascendante, et afin d'obtenir l'image segmentée en régions homogènes, nous adoptons deux stratégies de raffinement-estimation des distributions de possibilités caractérisant les M classes (prédéfinies par l'expert dans la scène), comme schématisé en Figure 6.2 :

- La première stratégie consiste à réaliser le raffinement-estimation des distributions de possibilités pour les M classes et d'une façon individuelle pour chaque image de primitive sélectionnée.
- Dans la deuxième stratégie, le raffinement-estimation conjoint des distributions de possibilités (i.e. en prenant en compte les N images de primitives) pour les M classes est réalisé. Une valeur de possibilité conjointe $\pi_{P_0}(C_m, \{I_1, I_2, \dots, I_N\})$ représente dans quel mesure il est possible que le pixel P_0 soit simultanément un germe possibiliste de la classe C_m dans chacune des images de primitives $I_n, n=1, \dots, N$. La règle que nous proposons pour la prise d'une décision qu'un pixel considéré soit un germe possibiliste ou non, est basée sur les deux niveaux pixelique et spatial à la fois et pour chaque primitive (Equation (6.1)) :

$$P_0 \in \{I_1, I_2, \dots, I_N\} \text{ est un germe possibiliste ssi :}$$

$$\exists C_{m_0} \in \psi / \begin{cases} \text{Ind}(\{C_{m_0}, I_n\}) \geq S, \\ \widetilde{\text{Ind}}(\{C_{m_0}, I_n\}) = \widetilde{\pi}_{P_0}(C_{m_0}, I_n) - \max_{m \neq m_0} \widetilde{\pi}_{P_0}(C_m, I_n) \geq S \end{cases} \quad n=1, \dots, N \quad (6.1)$$

où $\text{Ind}(\{C_{m_0}, I_n\})$ est l'indice de confiance de la classe C_{m_0} au niveau pixelique et en prenant en compte l'image de primitive I_n et S est le seuil de confiance décisionnelle appliqué sur les deux niveaux : pixelique et spatial. Notons que le raffinement-estimation conjoint des distributions de

possibilités est effectué dans le but d'augmenter la certitude dans le processus d'extraction des germes possibilistes.

L'étape de raffinement-estimation des distributions de possibilités, pour les deux stratégies adoptées, image/image et conjoint, aboutit à la construction de N distributions de possibilités pour chacune des M classes considérées dans la scène analysée. Cette étape de raffinement est suivie par une étape de projection des connaissances possibilistes permettant d'avoir N cartes possibilistes pour chaque classe, notées $CP_{I_n, C_m}(H \times W)$, $m=1, \dots, M$ et $n=1, \dots, N$.

Pour chaque classe C_m , les N cartes possibilistes CP_{I_n, C_m} sont ensuite fusionnées en utilisant l'un des opérateurs de fusion proposés dans le cadre possibiliste. Etant donné un pixel P, le degré de possibilité pour que ce pixel soit de la classe C_m est donné par :

$$\begin{aligned} CP_{C_m}(i,j) &= CP_{I_1, C_m}(x_1(P(i,j))) * \dots * CP_{I_N, C_m}(x_N(P(i,j))) \\ &= \pi_{I_1, C_m}(x_1(P(i,j))) * \dots * \pi_{I_N, C_m}(x_N(P(i,j))), \quad i=1, \dots, H; j=1, \dots, W \end{aligned}$$

où * est un opérateur de fusion, x_n , $n=1, \dots, N$ désigne la primitive observée au pixel P(i,j) et $\pi_{I_n, C_m}(x_1(P(i,j)))$ est le degré de possibilité d'être de la classe C_m étant donné la primitive n.

Cette fusion aboutit à avoir M cartes possibilistes représentant chacune des M classes. En exploitant ces cartes possibilistes, la diffusion croisée des connaissances possibilistes est réalisée. La maximisation de l'indice de confiance *Ind* est utilisée comme règle de décision et ceci afin de délimiter les régions homogènes. Rappelons que cet indice représente l'écart entre les deux plus grands degrés de possibilité sur l'univers de discours. Nous avons adopté la version la plus stricte de cette règle de décision (Equation (2.18)) qui n'accepte la prise de décision que si la valeur de l'indice de confiance *Ind* dépasse un certain seuil S prédéfini. Notons qu'une valeur élevée de ce seuil permet d'avoir une idée rapide des régions homogènes « pures » dans la scène analysée.

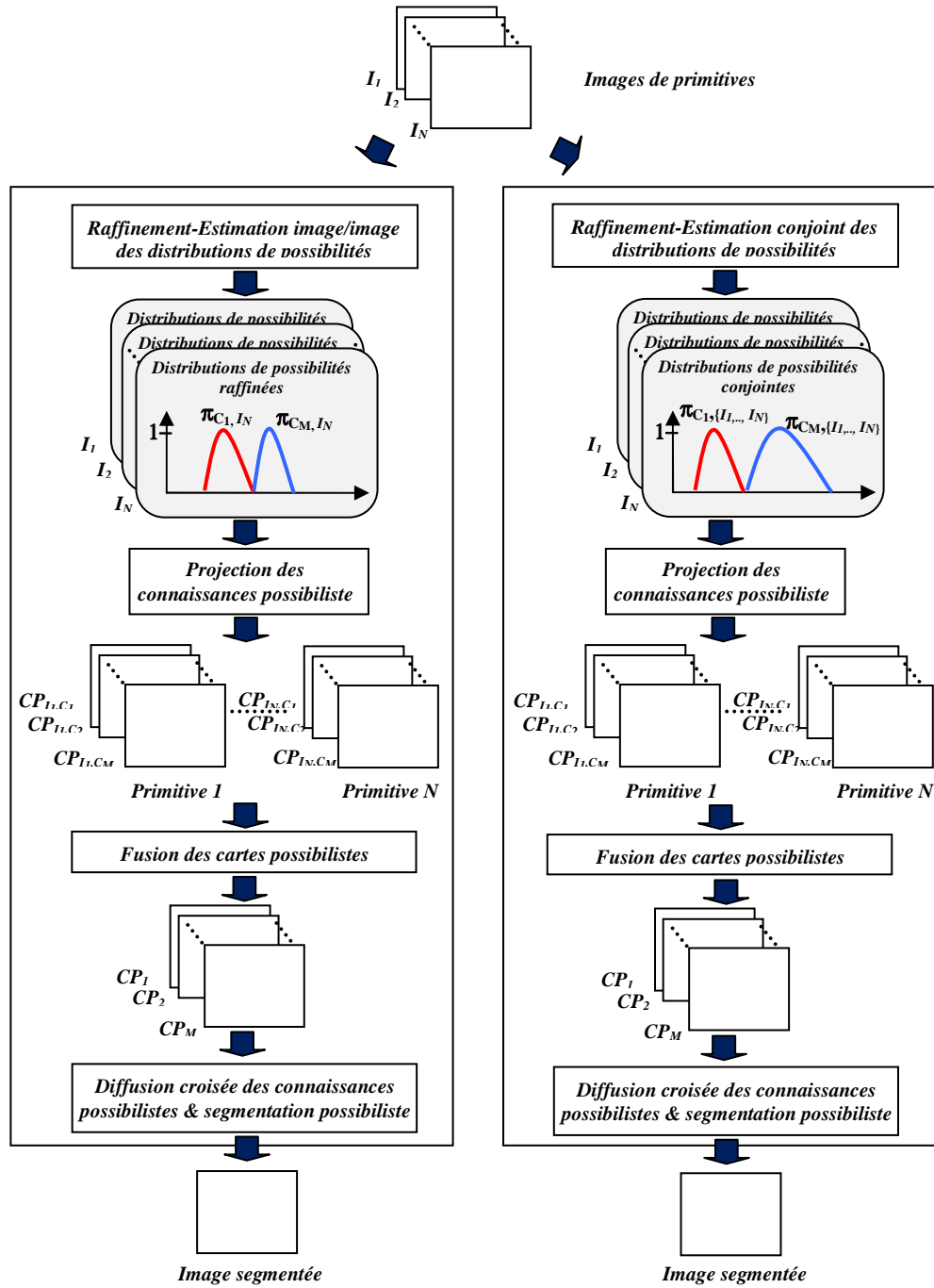


Figure 6.2 : Phase ascendante du système proposé d'interprétation de scène et les deux stratégies de raffinement-estimation des distributions de possibilités adoptées : le raffinement-estimation des distributions de possibilités d'une façon individuelle (gauche) et le raffinement-estimation conjoint des distributions de possibilités (droite).

6.3 Application de la phase ascendante à une image de télédétection

Dans ce paragraphe, nous proposons illustrer les résultats obtenus par l'application de la phase ascendante du système d'interprétation de scène proposée sur un exemple en télédétection. Dans cet exemple, la scène qui contient quatre classes thématiques, est observée par le capteur multi-spectral SPOT sur trois bandes spectrales qui sont montrées en Figure 6.3 [148].

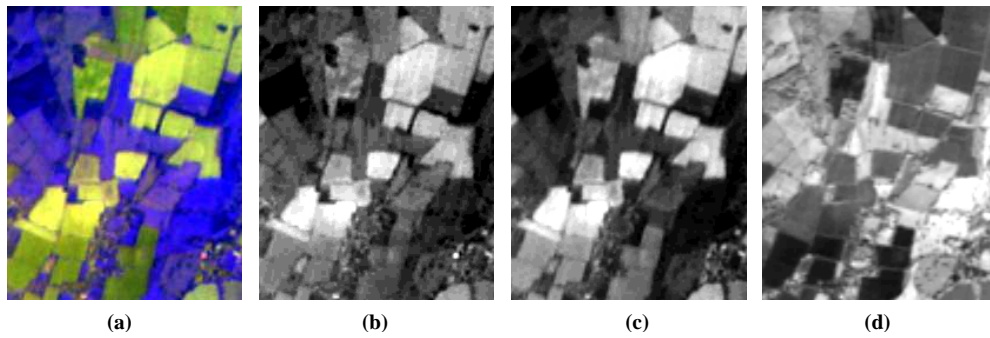


Figure 6.3 :(a) Image SPOT composée de quatre classes, (b) première bande, (c) deuxième bande, et (d) troisième bande

Quatre zones d'apprentissages de taille 10×10 pixels correspondant aux quatre classes sont délimitées par un expert, sur l'image SPOT. Ces zones sont indiquées en Figure 6.4.

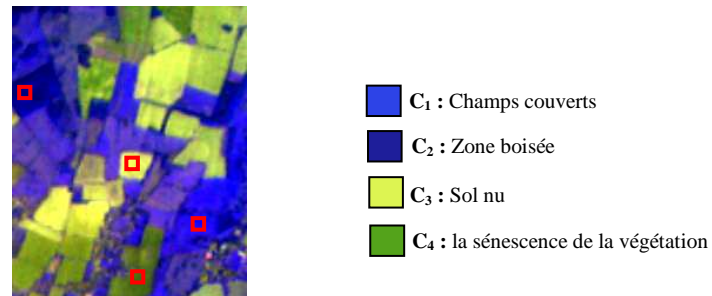


Figure 6.4 : Image du capteur SPOT et zones d'apprentissage considérées

6.3.1 Le raffinement-estimation des distributions de possibilités

Dans cette phase, et dans un premier temps, nous appliquons le processus de raffinement-estimation image/image, pour chaque bande spectrale de l'image SPOT). Dans un deuxième temps, nous appliquons le processus de raffinement-estimation conjoint. Ceci, pour les deux stratégies considérées, nous permet d'obtenir quatre distributions de possibilités raffinées caractérisant les quatre classes prédéfinies dans l'image SPOT, pour chacune des images de primitives (Figure 6.5, Figure 6.6, et Figure 6.7).

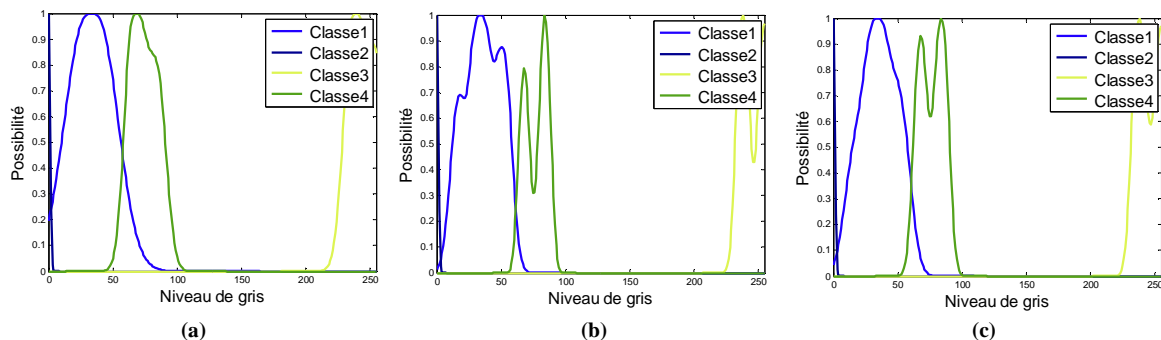


Figure 6.5 : Distributions de possibilités pour les quatre classes dans la première bande (a) distributions de possibilités initiales, (b) distributions de possibilités raffinées image/image et (c) distributions de possibilités raffinées conjointement

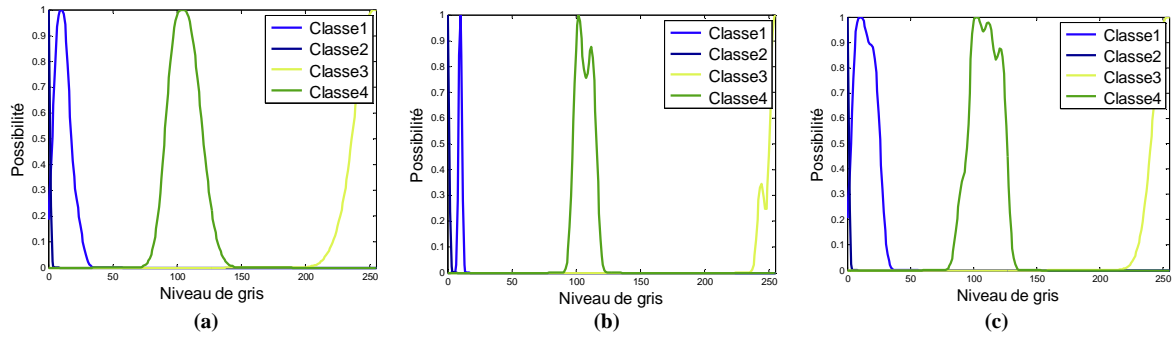


Figure 6.6 : Distributions de possibilités pour les quatre classes dans la deuxième bande (a) distributions de possibilités initiales, (b) distributions de possibilités raffinées image/image et (c) distributions de possibilités raffinées conjointement

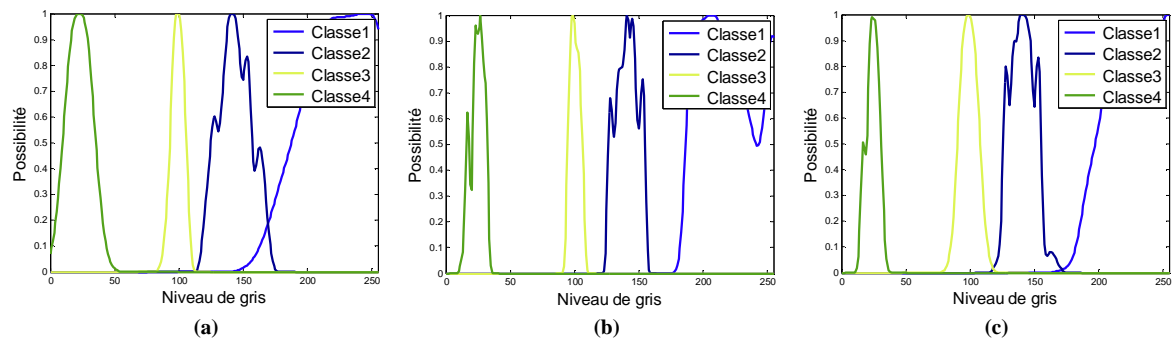


Figure 6.7 : Distributions de possibilités pour les quatre classes dans la troisième bande (a) distributions de possibilités initiales, (b) distributions de possibilités raffinées image/image et (c) distributions de possibilités raffinées conjointement

Nous remarquons l'aspect plus « lisse » des distributions de possibilités raffinées dans le cas du raffinement-estimation conjoint par rapport au cas du raffinement estimation image/image, ce qui est dû aux contraintes imposées pour sélectionner les germes possibilistes dans le cas du raffinement-estimation conjoint.

6.3.2 La projection des connaissances possibilistes

L'étape du raffinement-estimation des distributions de possibilités aboutit, dans chaque bande spectrale I_n , $n=1, 2, 3$, à la construction de quatre distributions de possibilités (associées aux quatre classes thématiques $\psi = \{C_1, C_2, C_3, C_4\}$ présentes dans la scène analysée et prédéfinies par l'expert dans les images observées I_1, I_2, I_3). Dans chaque bande spectrale, l'application de ces quatre distributions de possibilités sur l'image associée permet d'avoir quatre cartes possibilistes: CP_{I_n, C_m} , $n=1, 2, 3$ et $m=1, 2, 3, 4$ correspondant aux degrés de possibilités des différents pixels aux quatre classes prédéfinies dans la scène analysée.

Après la phase de raffinement-estimation des distributions de possibilités ainsi que celle de la projection de ces connaissances possibilistes, une phase de fusion des cartes possibilistes est effectuée. Cette phase de fusion aboutit à une carte possibiliste unique pour chacune des quatre classes prédéfinies : CP_{C_m} , $m=1, 2, 3, 4$.

Etant donné que les bandes spectrales peuvent être considérées comme des sources « concordantes » fournissant des informations qui sont en accord, nous adoptons, pour la fusion des cartes possibilistes, l'opérateur conjonctif « min » (Zadeh T-norme) [73]:

$$\begin{aligned}
CP_{C_m}(i,j) &= CP_{I_1,C_m}(x_1(P(i,j))) * \dots * CP_{I_N,C_m}(x_N(P(i,j))) \\
&= \min(\pi_{I_1,C_m}(x_1(P(i,j))), \dots, \pi_{I_N,C_m}(x_N(P(i,j))))
\end{aligned}$$

Les résultats de la fusion des cartes possibilistes initiales, raffinement-estimation image/image et raffinement-estimation conjoint des distributions de possibilités sont présentés dans les Figure 6.8, Figure 6.9 et Figure 6.10.

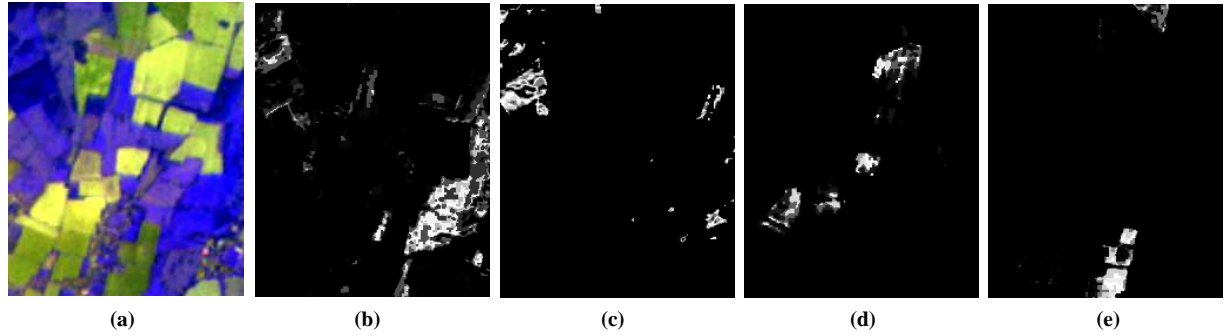


Figure 6.8 : Cartes possibilistes pour les quatre classes après le processus de la fusion (sans raffinement des distributions de possibilités), (a) image originale, (b) carte possibiliste de la classe C_1 , (c) carte possibiliste de la classe C_2 , (d) carte possibiliste de la classe C_3 , et (e) carte possibiliste de la classe C_4

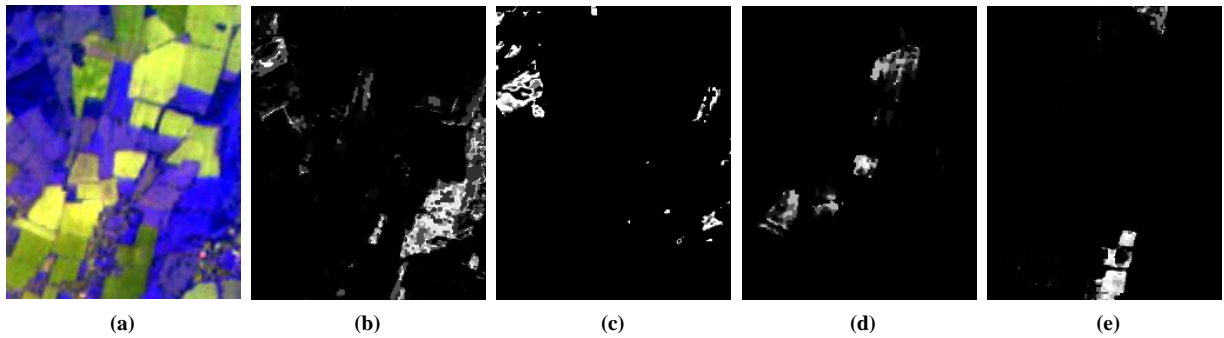


Figure 6.9 : Cartes possibilistes pour les quatre classes après le processus de la fusion (après raffinement-estimation image/image des distributions de possibilités), (a) image originale, (b) carte possibiliste de la classe C_1 , (c) carte possibiliste de la classe C_2 , (d) carte possibiliste de la classe C_3 , et (e) carte possibiliste de la classe C_4

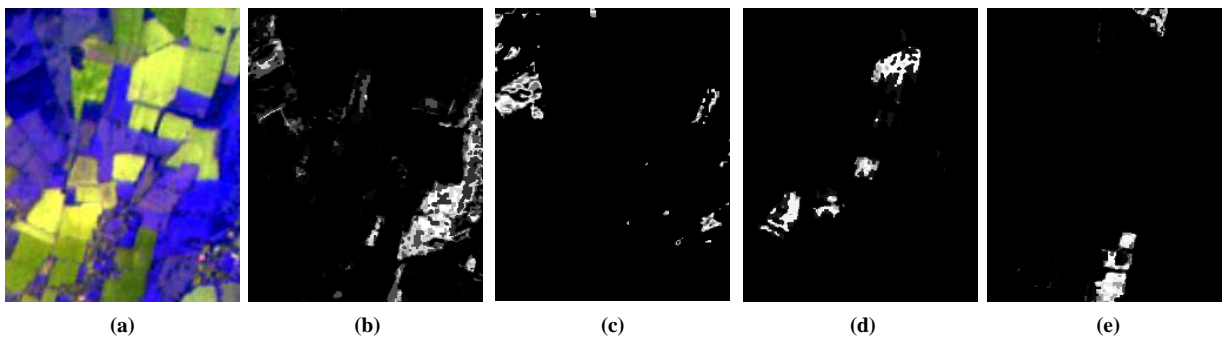


Figure 6.10 : Cartes possibilistes pour les quatre classes après le processus de la fusion (après raffinement-estimation conjoint des distributions de possibilités), (a) image originale, (b) carte possibiliste de la classe C_1 , (c) carte possibiliste de la classe C_2 , (d) carte possibiliste de la classe C_3 , et (e) carte possibiliste de la classe C_4

Il est bien évident qu'une analyse visuelle des résultats ne permet pas de donner une préférence de l'approche à adopter (parmi les trois présentées : sans raffinement/estimation, avec raffinement/estimation image par image et, avec raffinement/estimation conjointe). En effet, les trois approches

montrent des résultats similaires. Une légère préférence (visuelle) est attribuée à l'approche avec raffinement/estimation conjointe car nous remarquons une homogénéité plus marquée des zones de forts degrés de possibilités pour les différentes classes. De plus, notons la présence de grandes zones de faibles degrés de possibilités aux différentes classes, ceci signifie qu'il s'agit de grandes zones mixtes.

6.3.3 La diffusion des connaissances possibilistes et segmentation possibiliste

Le processus de segmentation est accompagné d'un processus de classification dans lequel nous proposons d'ajouter une cinquième classe représentant les régions homogènes « mixtes » qui sont difficiles à segmenter (on suppose que nous opérons sous l'hypothèse du monde fermé, c'est-à-dire seules les quatre classes identifiées par l'expert sont présentes dans la scène analysée).

La règle de classification pixelique que nous adoptons se base sur la valeur de l'indice de confiance *Ind*. En effet, si l'indice de confiance est supérieur à une valeur de seuil S_0 (i.e. il existe une classe unique ayant un degré de possibilité supérieur d'au moins S_0 à tous les autres degrés de possibilités), alors le pixel est attribué à cette classe. En revanche, si la valeur de l'indice de confiance est inférieure au seuil S_0 , alors le pixel est attribué à la classe C_5 qui est une classe de rejet de classification temporaire. Afin d'obtenir l'image segmentée en régions homogènes, la diffusion croisée des connaissances possibilistes est appliquée sur les cartes possibilistes issues de la phase de fusion (et ceci pour les trois cas des distributions de possibilités : initiales, raffinement-estimation image/image et raffinement-estimation conjoint).

Les résultats du processus de segmentation/classification adopté (pour une valeur de seuil d'indice de confiance $S_0 = 0$) sont présentés dans les Figure 6.11, Figure 6.12, et Figure 6.13.

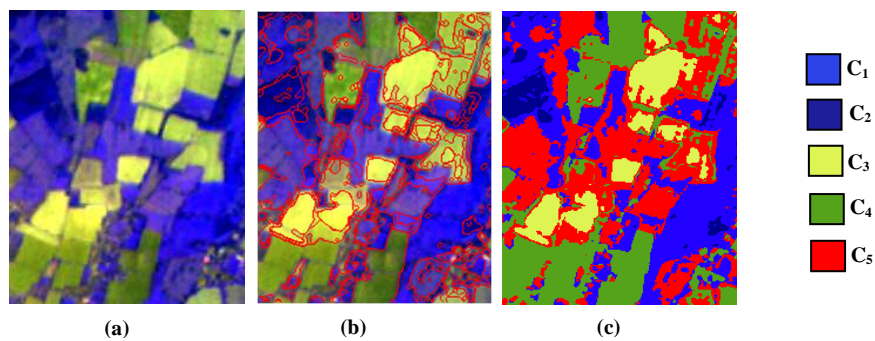


Figure 6.11 : Segmentation/classification de l'image SPOT sans raffinement des distributions de possibilités (a) image originale, (b) contours des régions identifiées dans l'image segmentée, (c) image classifiée

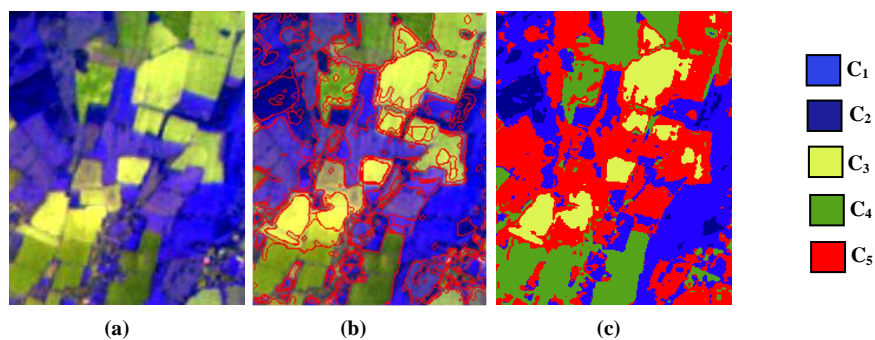


Figure 6.12 : Segmentation/classification de l'image SPOT au cas du raffinement-estimation image/image (a) image originale, (b) contours des régions identifiées dans image segmentée, (c) image classifiée

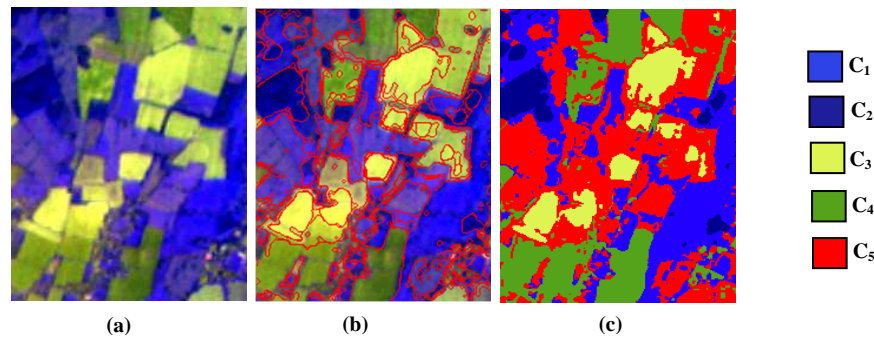


Figure 6.13 : Segmentation/classification de l'image SPOT au cas du raffinement-estimation conjoint (a) image originale, (b) contours des régions identifiées dans image segmentée, (c) image classifiée

Nous constatons pour les deux cas de raffinement image/image et raffinement conjoint, l'effet de raffinement en termes d'élimination des petites régions est plus visible permettant ainsi de mieux identifier les zones homogènes (à titre d'exemple, la zone mixte en haut à droite de la scène analysée).

De plus, le raffinement/estimation conjoint permet d'aboutir à des régions plus homogènes que les deux autres approches (en termes de régions détectées visuellement dans cette image).

Afin de montrer l'intérêt de l'indice de confiance pour l'inspection des régions homogènes « pures » ainsi que les régions homogènes « mixtes », nous avons utilisé, dans le cas de raffinement-estimation conjoint des distributions de possibilités, trois valeurs du seuil S_0 , $S_0 = 0.5$, 0.1 , et 0 . Les résultats de la classification pour ces trois valeurs de seuil d'indice de confiance sont présentés dans la Figure 6.14.

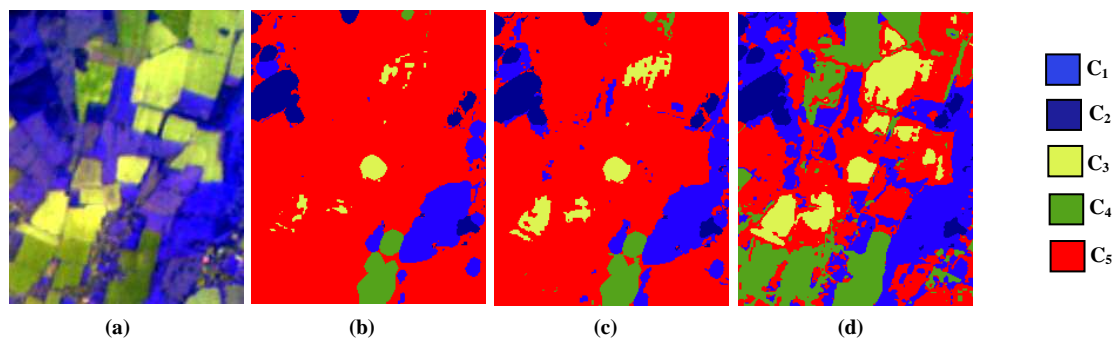


Figure 6.14 : Classification de l'image SPOT au cas du raffinement-estimation conjoint (a) image originale, (b) image classifiée pour $S_0 = 0.5$, (c) image classifiée pour $S_0 = 0.1$ et (d) image classifiée pour $S_0 = 0$

En effet, plus cette valeur de seuil est grande, plus les régions détectées sont pures. Nous pouvons constater que pour la valeur $S_0 = 0.5$, la classification de l'image SPOT en se basant sur la maximisation de l'indice de confiance Ind , nous permet d'avoir une idée sur la localisation des régions homogènes « pures » pour lesquelles on est totalement certains de leurs contenus sémantique dans la scène analysée. De plus, cet indice pour la valeur $S_0 = 0$, nous permet de localiser les régions homogènes « mixtes » comportant un mélange très important de plusieurs classes. Afin de mettre en évidence ces régions homogènes « mixtes », nous allons utiliser dans la suite, la valeur $S_0 = 0$ et ceci dans le cas de raffinement-estimation conjoint des distributions de possibilités.

6.4 Phase descendante et analyse de scène en télédétection

En élargissant l'approche descendante du système d'interprétation de scène proposé, une étape de modélisation des régions est effectuée. Cette étape a pour objectif de constituer un formalisme permettant d'incorporer des connaissances *a priori* pour chacune des régions détectées dans la phase ascendante. Ce formalisme est composé des rapports où chaque rapport concerne une région détectée et comporte les

informations caractéristiques de cette région telles que le vecteur descriptif, la localisation, la taille,... ainsi que des connaissances *a priori* liées à la taille, forme (par exemple, une forme polygonale pour les champs agricoles), orientation et contexte spatial des régions (ou objets) dans la scène. Ces connaissances *a priori* enregistrées dans les rapports constituent des contraintes imposées qui seront utilisées pour guider l'analyse de la scène. Des tels rapports doivent répondre à des questions du type : « quelles sont les régions adjacentes de la région d'intérêt ? ». L'exploitation des rapports des régions détectées permet de réaliser l'analyse de la scène et de fusionner les régions (i.e. *Region Merging*) selon des critères prédéfinis.

6.4.1 Modélisation des régions

La phase ascendante d'analyse de scène vise à réaliser le processus de segmentation/classification dans l'objectif d'extraire, de labéliser et de calculer les paramètres géométriques (taille, forme, position, localisation des régions adjacentes etc.) des régions identifiées dans la scène analysée (Figure 6.15).

Dans la phase descendante, chaque région identifiée est analysée plus finement (en appliquant l'approche de démixage pixelique possibiliste proposée) dans l'objectif d'extraire le vecteur descriptif décrivant la distribution statistique des taux d'abondance des pixels composant cette région, aux différentes classes thématiques.

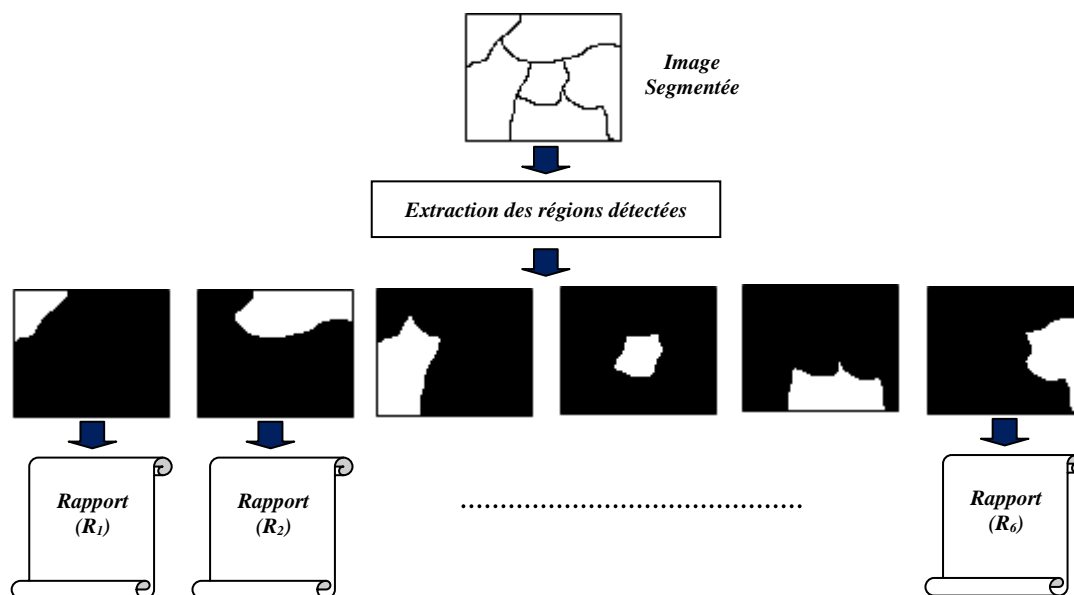


Figure 6.15 : Système de modélisation des régions

Ainsi, les informations essentielles caractérisant chaque région identifiée sont extraites et enregistrées dans le rapport descriptif lié à cette région. De plus, lors de la phase d'analyse ascendante, l'expert peut imposer une taille minimale des régions identifiées (afin d'éviter d'avoir un nombre élevé de régions de faibles tailles). A titre d'exemple, la Figure 6.16 illustre les résultats de la phase d'analyse ascendante d'une image SPOT segmentée par raffinement-estimation conjointe des distributions de possibilités suivi par la diffusion des connaissances possibilistes (sous la contrainte d'avoir des régions de taille minimale $T_m=50$ pixels). Les régions de faible taille sont supprimées en labélisant leurs pixels avec le même label de classe que la région englobant la région de faible taille.

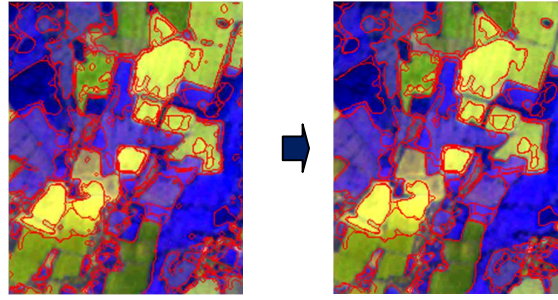


Figure 6.16 : Application de la contrainte liée à la taille des régions détectées (a) image segmentée par raffinement-estimation conjoint et avant l'application de la contrainte $T_m=50$ pixels (b) image segmentée par raffinement-estimation conjoint et après l'application de la contrainte $T_m=50$ pixels

On constate clairement que l'application d'une telle contrainte, permet de supprimer les petites régions qui sont souvent des fausses alarmes. A l'issue de la phase ascendante, l'ensemble des paramètres géométriques concernant chaque région identifiée, est calculé et enregistré dans le rapport descriptif de cette région. La phase descendante consiste à appliquer l'approche de démixage pixelique possibiliste sur chacune des régions identifiées dans la phase ascendante.

Au niveau de chaque source d'information (i.e. bande spectrale dans ce cas) les distributions de possibilités initiales (estimées à partir des zones d'apprentissage) permettent d'obtenir les cartes d'abondance pour l'ensemble des classes thématiques. Au niveau de chaque classe thématique, la carte d'abondance finale est obtenue comme étant la moyenne des cartes d'abondance des différentes sources d'information. Les résultats obtenus par l'application de l'approche proposée sur l'image SOPT, en termes de cartes d'abondance, sont donnés dans la Figure 6.17.

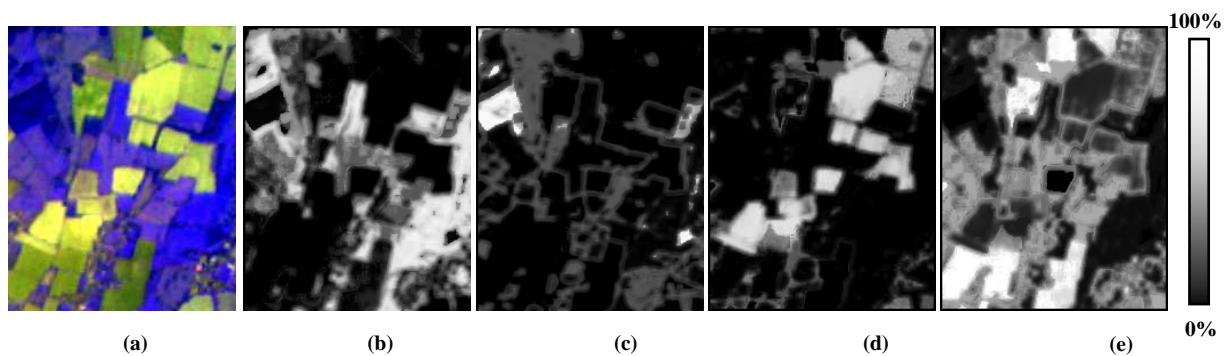


Figure 6.17 : Cartes d'abondance pour les quatre classes thématiques (a) image originale (b) carte d'abondance de la classe C_1 (c) carte d'abondance de la classe C_2 (d) carte d'abondance de la classe C_3 et (e) carte d'abondance de la classe C_4

Une analyse visuelle de ces cartes montre un chevauchement entre d'un côté les classes C_1 (Champs couverts), C_2 (Zone boisée) et C_4 (la sénescence de la végétation), et de l'autre côté entre la classe C_3 (Sol nu) et la classe C_4 (la sénescence de la végétation). En revanche, peu de chevauchement est constaté entre la classe C_3 (Sol nu) et les deux autres classes C_1 et C_2 .

Pour chaque région identifiée, le vecteur de descripteurs comporte les quatre histogrammes représentant les distributions statistiques des taux d'abondance des différentes classes dans cette région.

Les Figure 6.18, Figure 6.19 et Figure 6.20 illustrent des exemples de régions identifiées par le biais de la phase ascendante sur l'image SPOT considérée ainsi que les vecteurs descriptifs associés à ces

régions. La Figure 6.18 présente une région « pure » (i.e. ne comportant qu'une seule classe thématique) tandis que les deux autres figures : Figure 6.19 et la Figure 6.20 présentent des régions « mixtes ».

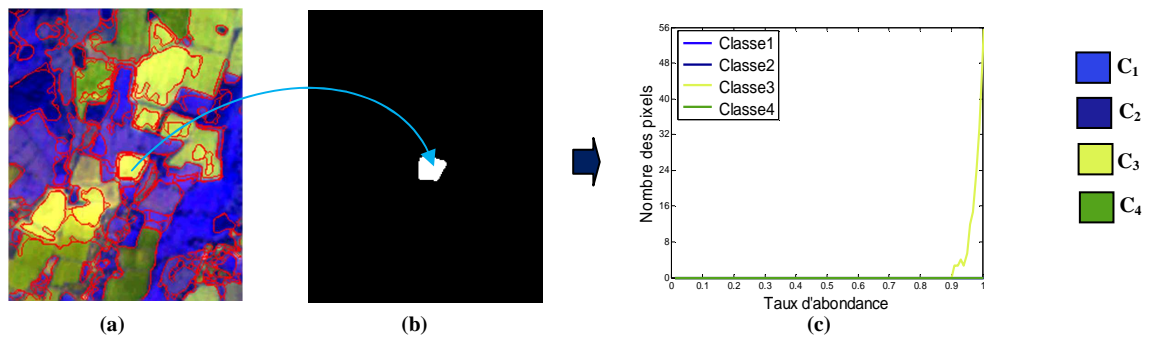


Figure 6.18 : Exemple d'une région homogène-pure sur l'image SPOT : (a) Image originale segmentée sous la contrainte $T_m=50$ pixels, (b) La région homogène- pure, et (c) Le vecteur descriptif associé

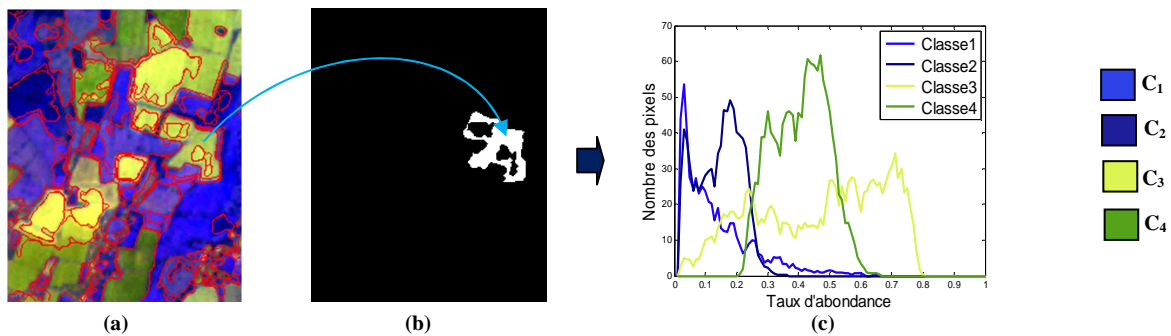


Figure 6.19 : Exemple d'une région homogène-mixte comportant un fort mélange des quatre classes thématiques : (a) Image originale segmentée sous la contrainte $T_m=50$ pixels, (b) La région homogène-mixte, et (c) Le vecteur descriptif associé

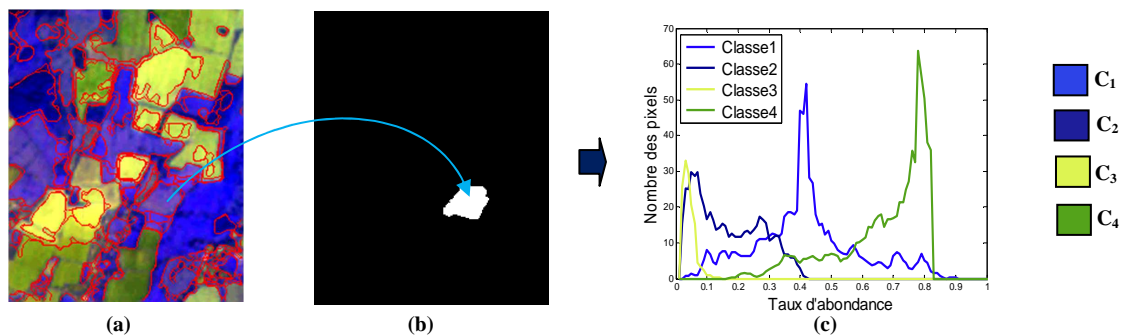


Figure 6.20 : Exemple d'une région homogène-mixte comportant un fort mélange des quatre classes thématiques : (a) Image originale segmentée sous la contrainte $T_m=50$ pixels, (b) La région homogène-mixte, et (c) Le vecteur descriptif associé

Ces exemples montrent l'un des intérêts fondamentaux des vecteurs descriptifs proposés par cette démarche. En effet, ils peuvent confirmer/infirmier l'aspect pure d'une région identifiée et donner une information essentielle sur le contenu qualitatif (i.e. les classes thématiques contenues) et quantitatifs (i.e. les taux d'abondance des classes thématiques) des régions identifiées. Nous remarquons que

l'histogramme des taux d'abondance des classes C_3 dans la région illustrée dans la Figure 6.18, présente des valeurs d'abondance centrées sur la valeur « 1 », ce qui suggère fortement la pureté de la classe C_3 . De plus, les histogrammes des taux d'abondance des classes C_1 , C_2 , C_3 et C_4 dans les régions illustrées dans la Figure 6.19 et la Figure 6.20 permettent d'apprécier le mélange de ces classes. Notons, à titre d'exemple, que l'inspection visuelle de la région illustrée dans la Figure 6.19, laisse à penser qu'il s'agit d'une région mixte ne comportant que les contenus thématiques : C_3 et C_4 pourtant, nous remarquons la présence de taux d'abondance important des deux autres classes thématiques C_1 et surtout C_2 .

Les histogrammes des taux d'abondance des classes C_1 , C_2 , C_3 et C_4 dans les différentes régions identifiées dans la phase ascendante seront exploités dans la suite afin d'étudier la similarité entre ces régions en termes du contenu thématique. Une telle étude de similarité est menée dans le but d'accomplir l'interprétation de la scène (au-delà de l'inspection visuelle).

6.4.2 Similarité de régions

Une fois l'ensemble des régions contenues dans la scène analysée identifié, trois questions s'imposent:

1. Etant donnée une région spécifique dans la scène, quelles sont les régions qui lui sont similaires dans la scène analysée ?
2. Etant donnée une classe spécifique dans la scène, quelles sont les régions qui contiennent cette classe ?
3. *Region Merging* : Etant données deux régions adjacentes, à quel niveau sont-elles similaires au point de les fusionner en une seule région (ou sinon de bien les garder comme étant deux régions distinctes) ?

Ces trois questions impliquent la définition d'une mesure de similarité entre deux régions identifiées dans la scène analysée. Pour le calcul de cette mesure de similarité, nous allons nous contenter des vecteurs descriptifs comportant les taux d'abondance sans considérer les paramètres géométriques qui caractérisent les différentes régions.

Chaque région R_k est caractérisée par un vecteur descriptif $V_{R_k} = [A_{R_k, C_1}(x), A_{R_k, C_2}(x), \dots, A_{R_k, C_M}(x)]$ où $x \in [0,1]$ désigne le taux d'abondance et $A_{R_k, C_m}(x)$ désigne l'histogramme de taux d'abondance de la classe thématique C_m dans la région R_k .

Similarité liée à une classe thématique

Dans la littérature, des mesures de distance sont utilisées pour déterminer à quel point un histogramme diffère d'un autre histogramme (i.e. pour mesurer la « dissimilarité » entre les histogrammes)[149]. Etant donné que les histogrammes peuvent être considérés comme des vecteurs numériques, les mesures de distance métriques telles que les mesures de *Minkowski* (voir chapitre 4), peuvent être utilisées. D'autres techniques de comparaison d'histogrammes ont été introduites dans la littérature, nous citons, la distance χ^2 -test qui est basée sur les statistiques [150], l'intersection des histogrammes [151] et les techniques basées sur la comparaison des histogrammes cumulatifs [149]. Malgré cette diversité des mesures de similarité entre les histogrammes proposées dans la littérature, le choix d'une mesure convenable reste toujours une question ouverte.

De plus, par une simple normalisation des histogrammes (afin d'avoir une valeur maximale égale à l'unité) les histogrammes normalisés peuvent être considérés comme des fonctions d'appartenance à des ensembles flous. Par conséquent, les mesures de similarité, introduites à l'origine pour quantifier la ressemblance entre deux ensembles flous, peuvent aussi être appliquées pour la comparaison de deux histogrammes [149].

Pour la sélection de la mesure de similarité entre histogrammes, Van der Weken *et al.* [149] ont proposé quatre propriétés qu'une telle mesure doit vérifier :

1. **La réflexivité** : Deux histogrammes identiques doivent avoir une mesure de similarité unitaire ;
2. **La symétrie** : La mesure de similarité entre deux histogrammes est indépendante de l'ordre de leur présentation ;
3. **La résistance au bruit** : Une bonne mesure de similarité ne doit pas être trop influencée par le bruit, et doit être décroissante en fonction de l'augmentation du niveau de bruit ;
4. **La résistance au décalage des valeurs mesurées** : Une bonne mesure de similarité doit être performant et ne pas être trop affectée par un faible décalage des valeurs mesurées (phénomènes de « *Enlightening* » et « *Darkening* »). Cette mesure doit être aussi décroissante en fonction de l'augmentation du décalage des valeurs mesurées.

Une étude expérimentale de l'applicabilité des mesures de similarité aux histogrammes normalisés a abouti à plusieurs mesures de similarité données dans [149]. Parmi ces mesures de similarité d'histogrammes proposées, nous avons retenu la mesure de similarité suivante entre l'histogramme du taux d'abondance $A_{R_k, C_m}(x)$ et l'histogramme du taux d'abondance $A_{R_h, C_2}(x)$:

$$\begin{aligned} \text{Sim}_H(A_{R_k, C_m}, A_{R_h, C_m}) &= \frac{|A_{R_k, C_m}(x) \cap A_{R_h, C_m}(x)|}{|A_{R_k, C_m}(x) \cup A_{R_h, C_m}(x)|} \\ &= 1 \text{ si } A_{R_k, C_m} = A_{R_h, C_m} = 0 \end{aligned} \quad (6.2)$$

Cette mesure prend ses valeurs dans l'intervalle $[0,1]$ et les deux histogrammes normalisés A_{R_k, C_m} et A_{R_h, C_m} sont considérés comme étant des fonctions d'appartenance à des ensembles flous, \cap et \cup indiquent les opérations d'intersection et d'union des ensembles flous et $|A|$ désigne la cardinalité de l'ensemble A .

Notons que la majorité des mesures proposées dans [149] apportent des résultats comparables. La raison du choix de cette mesure Sim_H réside dans sa capacité, d'un côté, de quantifier la zone de chevauchement entre les histogrammes et, d'autre côté, d'apporter une solution au cas trivial où l'un de deux histogrammes comparés est nul (i.e. la région ne contient pas la classe correspondante).

A titre d'exemple, la Figure 6.21 présente cinq exemples de régions R_1 , R_2 , R_3 , R_4 et R_5 identifiées en appliquant la phase ascendante sur l'image SPOT et les vecteurs descriptifs associés à ces régions.

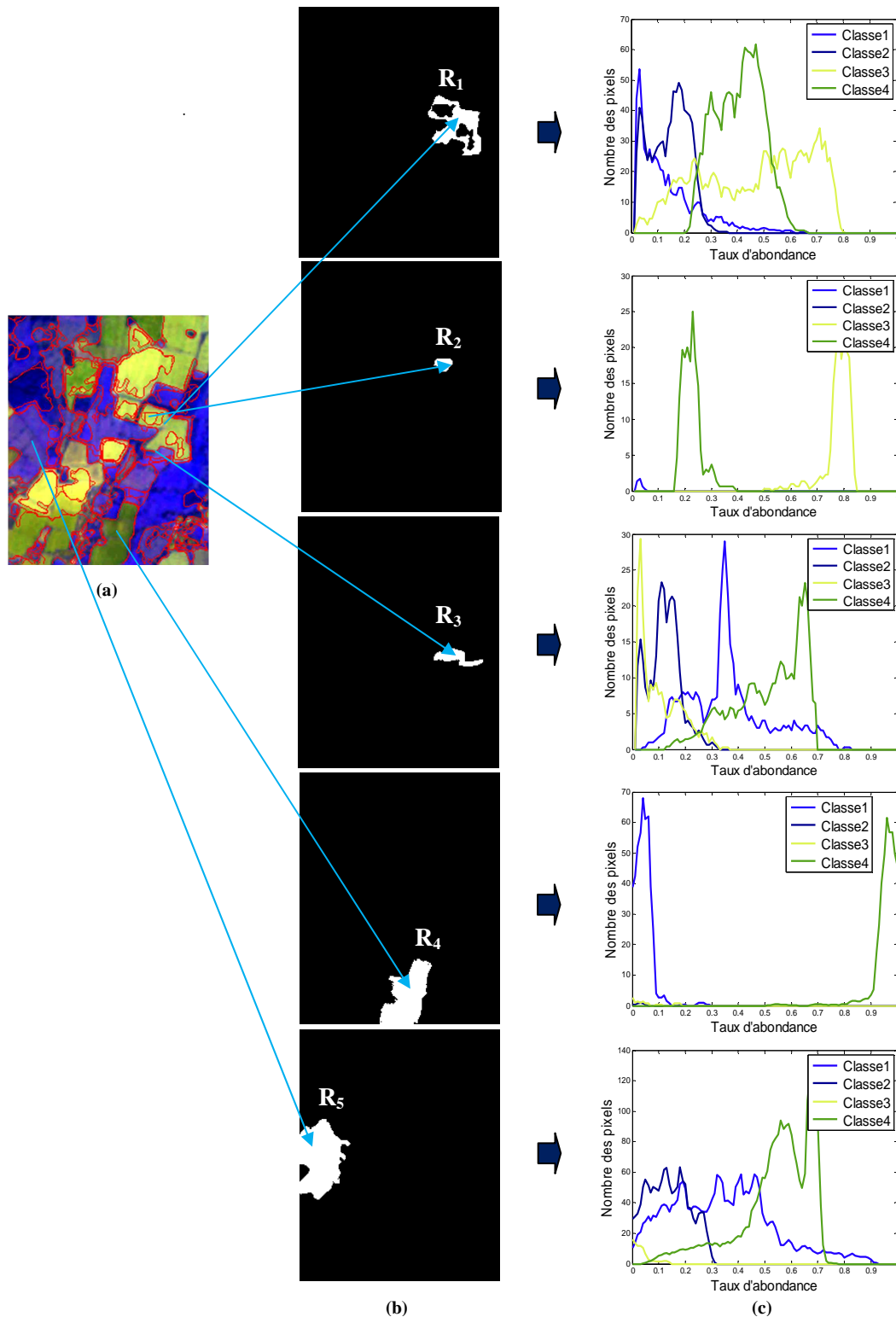


Figure 6.21 : Exemple de régions identifiées en appliquant la phase ascendante et leurs vecteurs descriptifs (a) image segmentée, (b) Régions R₁, R₂, R₃, R₄ et R₅ et (c) vecteurs descriptifs

Notons que les trois régions R₁, R₂ et R₃ sont adjacentes et sont donc potentiellement candidates à une fusion (i.e. *réunion merging*) et ceci malgré la différence visuelle entre les deux régions R₁ et R₃. La sélection de deux régions « éloignées » R₄ et R₅ a pour objectif d'étudier la similarité en termes de distribution de classes dans la scène.

Concernant les deux régions R₁ et R₂ (R₂ est encadrée par la région R₁), et en considérant la mesure de similarité choisie Sim_H (Equation (6.2)), nous obtenons les résultats suivants :

- $\text{Sim}_H(A_{R_1,C_1}, A_{R_2,C_1}) = 0.4$;
- $\text{Sim}_H(A_{R_1,C_2}, A_{R_2,C_2}) = 0$;
- $\text{Sim}_H(A_{R_1,C_3}, A_{R_2,C_3}) = 0.05$;
- $\text{Sim}_H(A_{R_1,C_4}, A_{R_2,C_4}) = 0.07$.

Notons :

- L'absence de similarité entre les régions R_1 et R_2 relativement à la classe C_2 (et ceci malgré la présence de cette classe dans la région R_1) signifie l'absence totale de la classe C_2 dans la composition des pixels de la région R_2 ;
- La similarité la plus élevée est celle liée à la classe C_1 et ceci malgré la faible présence de cette classe dans la région R_2 . En effet, les histogrammes d'abondance donnés dans la Figure 6.21 ne sont pas normalisés et la région R_2 est de faible taille comparée à la région R_1 . La normalisation des histogrammes permet de s'affranchir de la contrainte de taille et met en évidence la forte similarité des distributions d'abondance de la classe C_1 dans les deux régions (Figure 6.22) ;

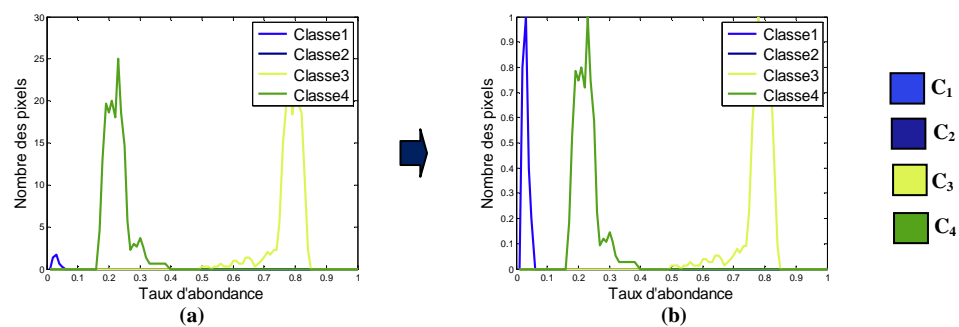


Figure 6.22 : Exemple de l'effet de la normalisation sur les histogrammes d'abondance liée aux classes dans la région R_2
(a) les histogrammes d'abondance non normalisés et (b) les histogrammes d'abondance normalisés

- L'analyse visuelle de la similarité entre les deux régions R_1 et R_2 est fortement influencée par la similarité des histogrammes d'abondances liés aux classes C_3 et C_4 malgré que les valeurs mesurées soient plus faible que celle concernant la classe C_2 . Notons que les faibles mesures de similarité entre les histogrammes d'abondances liés aux classes C_3 et C_4 sont dus de la faible ampleur des valeurs de taux d'abondance liés aux classes C_3 et C_4 dans la région R_2 (i.e. concentration des valeurs de taux d'abondance autour certaines valeurs)

Concernant les deux régions R_1 et R_3 qui sont adjacentes mais d'apparences non similaires, les mesures de similarités des histogrammes d'abondances sont les suivantes :

- $\text{Sim}_H(A_{R_1,C_1}, A_{R_3,C_1}) = 0.24$;
- $\text{Sim}_H(A_{R_1,C_2}, A_{R_3,C_2}) = 0.57$;
- $\text{Sim}_H(A_{R_1,C_3}, A_{R_3,C_3}) = 0.11$;
- $\text{Sim}_H(A_{R_1,C_4}, A_{R_3,C_4}) = 0.27$.

Malgré l'apparence visuelle de différence entre ces deux régions (affichées en format RGB), ces deux régions sont fortement similaires. En effet, les quatre classes contenues dans la scène sont présentes dans ces deux régions. Notons que la « couleur » dominante dans une région est fortement influencée par la forte abondance de certaines classes. Néanmoins, cette représentation en format RGB (classiquement utilisée en imagerie satellitaire) ne tient pas compte de la similarité des histogrammes d'abondance (même avec des faibles valeurs).

Analysons brièvement les mesures de similarité d'histogrammes d'abondance entre les deux couples de régions (R_1, R_3) et (R_3, R_5) :

- $\text{Sim}_H(A_{R_1,C_1}, A_{R_3,C_1}) = 0.24$; $\text{Sim}_H(A_{R_3,C_1}, A_{R_5,C_1}) = 0.37$;
- $\text{Sim}_H(A_{R_1,C_2}, A_{R_3,C_2}) = 0.57$; $\text{Sim}_H(A_{R_3,C_2}, A_{R_5,C_2}) = 0.49$;
- $\text{Sim}_H(A_{R_1,C_3}, A_{R_3,C_3}) = 0.11$; $\text{Sim}_H(A_{R_3,C_3}, A_{R_5,C_3}) = 0.56$;
- $\text{Sim}_H(A_{R_1,C_4}, A_{R_3,C_4}) = 0.27$. $\text{Sim}_H(A_{R_3,C_4}, A_{R_5,C_4}) = 0.66$.

En effet, les deux régions R_3 et R_5 semblent (visuellement parlant) être extrêmement similaires. Cette apparence de similarité visuelle est très fortement dominée par la dominance des abondances des classes C_1 et C_4 . Pourtant, les histogrammes d'abondance liés aux classes C_2 et C_3 sont aussi fortement similaires malgré leurs faibles abondances. De plus, les histogrammes d'abondance liés aux classes C_2 dans les (R_1, R_3) sont plus similaires que ceux entre les régions (R_3, R_5) et ceci malgré l'apparence visuelle similaire entre leurs les régions (R_3, R_5).

Cette analyse de similarité des histogrammes des taux d'abondance montre une facette importante de l'interprétation de scène qui ne se limite pas à la seule labellisation de région (i.e. classification globale). Elle permet de qualifier d'une façon plus fine le contenu thématique des différentes régions ce qui est partiellement ignoré dans une fonction de classification.

Similarité globale entre deux régions

Considérons deux régions R_k et R_h de la scène analysée. En supposant $\text{Sim}_H(A_{R_k,C_m}, A_{R_h,C_m})$ la mesure de similarité que nous considérons entre les histogrammes décrivant le taux d'abondance de la classe thématique C_m , $m=1, 2, \dots, M$ dans les deux régions considérées, alors la mesure de similarité globale $\text{Sim}_G(V_{R_k}, V_{R_h})$ entre ces deux régions R_k et R_h peut être définie par le biais d'une fonction \mathcal{F} permettant de combiner l'ensemble des mesures de similarités pour toutes les classes thématiques dans les deux régions:

$$\mathcal{F}: V \times V \rightarrow [0,1]$$

$$(V_{R_k}, V_{R_h}) \rightarrow \text{Sim}_G(V_{R_k}, V_{R_h}) = \mathcal{F}[\text{Sim}_H(A_{R_k,C_m}, A_{R_h,C_m})], \quad m = 1, \dots, M$$

où V représente l'ensemble des vecteurs descriptifs caractérisant les différentes régions identifiées dans la phase ascendante.

Dans le processus de construction des rapports descriptifs, les valeurs de similarité calculées ainsi que des informations liées à la localisation des régions (ainsi que les régions adjacentes) sont enregistrées dans le rapport descriptif de chaque région. Ces valeurs de similarité enregistrées sont exploitées par une fonction \mathcal{F} pour définir la similarité globale entre les régions. Cette similarité est à son tour exploitée dans la phase d'analyse de la scène afin d'une part d'effectuer une étape de fusion des régions détectées et d'autre part de fournir une analyse globale de la scène.

A ce stade, étant donné que les valeurs de similarité calculées sont issues des sources d'information indépendantes (i.e. les histogrammes des taux d'abondance), et afin de combiner ces valeurs de la similarité, tout en prenant en compte la complémentarité de ces sources d'informations, nous proposons l'application de la fonction \mathcal{F} qui n'est qu'un simple moyennage des valeurs de similarité liées aux différentes classes thématiques :

$$\text{Sim}_G(V_{R_k}, V_{R_h}) = \mathcal{F}[\text{Sim}_H(A_{R_k, C_m}, A_{R_h, C_m})] = \sum_{m=1}^M \frac{\text{Sim}_H(A_{R_k, C_m}, A_{R_h, C_m})}{M}$$

En considérant cette fonction \mathcal{F} , nous obtenons les résultats suivants (concernant les régions de la Figure 6.21) :

- Pour les deux régions (R_1, R_2) : $\text{Sim}_G(V_{R_1}, V_{R_2}) = 0.13$
- Pour les deux régions (R_1, R_3) : $\text{Sim}_G(V_{R_1}, V_{R_3}) = 0.32$
- Pour les deux régions (R_1, R_4) : $\text{Sim}_G(V_{R_1}, V_{R_4}) = 0.33$
- Pour les deux régions (R_1, R_5) : $\text{Sim}_G(V_{R_1}, V_{R_5}) = 0.26$
- Pour les deux régions (R_3, R_5) : $\text{Sim}_G(V_{R_3}, V_{R_5}) = 0.52$

Les valeurs obtenues pour la similarité entre les régions (R_1, R_2) et (R_1, R_3) soulèvent la question importante concernant les critères à appliquer pour la fusion de régions. En effet, si l'objectif de la fusion concerne le regroupement des régions en fonction des forts taux d'abondance des classes, alors tous les algorithmes de fusion de régions auront pour résultat de fusionner les deux régions R_1 et R_2 , $\text{Sim}_G(V_{R_1}, V_{R_2}) = 0.13$ plutôt que les deux régions R_1 et R_3 , $\text{Sim}_G(V_{R_1}, V_{R_3}) = 0.32$. En revanche, le fait de considérer la similarité en termes de similarités des histogrammes des taux d'abondance, l'approche que nous proposons favorisera la fusion entre (R_1, R_3) plutôt que (R_1, R_2) .

Notons que des résultats similaires sont obtenus en adoptant les autres mesures de similarité d'histogrammes. Par conséquent, il s'agit de deux logiques de raisonnement pour la fusion de régions qui peuvent avoir une certaine divergence en termes des résultats.

6.4.3 Rapport d'interprétation de scène

A l'issue de l'application de l'ensemble des méthodes de traitement et d'analyse de scène, un rapport récapitulatif est délivré. En général, ce rapport est donné sous la forme d'une image thématique dans laquelle les différentes régions détectées sont visualisées et des couleurs, prédéfinies par l'expert, indiquent les labels de ces régions. En réalité, il s'agit d'un rapport de classification de scène et non d'un rapport d'interprétation de scène. La question du contenu d'un rapport d'interprétation de scène est en réalité une question encore ouverte et dépend fortement de l'application considérée et des exigences et contraintes imposées par l'expert.

En effet, si l'objectif visé par l'expert consiste à étudier la dispersion spatiale des taux d'abondance des différentes classes thématiques dans une scène, la question de la classification et des régions identifiées devient secondaire. Dans ce cas, le rapport d'interprétation de scène sera réduit à un ensemble d'images des taux d'abondance pixelique avec autant d'images que le nombre de classes thématiques. Dans d'autres exemples d'application, certaines contraintes peuvent coupler le contenu thématique avec la forme géométrique des régions détectées (ce qui est le cas en imagerie satellitaire lorsque l'objectif est de détecter des régions/bâtiments de formes et de contenus spécifiques). En imagerie médicale, certaines applications considèrent qu'un rapport d'interprétation doit simplement indiquer les régions de tumeurs ainsi que des régions « suspectées ».

Cette diversité de formes de rapports d'interprétation, montre la grande difficulté pour la définition d'un rapport générique d'interprétation de scène. Il est bien évident, que l'objectif de notre travail ne peut pas couvrir cette notion de rapport d'interprétation d'une façon globale. Néanmoins, l'ensemble des

outils d'analyse et d'interprétation que nous avons proposé donne lieu à un ensemble important de moyens permettant de caractériser l'ensemble des pixels et des régions identifiées dans la scène.

Dans ce paragraphe, nous proposons d'illustrer un exemple de rapport d'interprétation de scène. Le processus de construction de ce rapport comporte essentiellement deux étapes :

- La fouille de connaissances liées à l'homogénéité des régions identifiées dans l'approche ascendante. Cette fouille est réalisée grâce aux vecteurs descriptifs des régions calculés dans la démarche descendante et permet de donner les centres d'inertie des taux d'abondance de chacune des classes thématiques dans les différentes régions identifiées. Le centre d'inertie des taux d'abondance pour une classe C_m dans une région R_k est défini par :

$$CI_{R_k, C_m} = \frac{\sum_{x \in [0,1]} A_{R_k, C_m}(x) \cdot x}{\sum_{x \in [0,1]} A_{R_k, C_m}(x)}$$

- L'étude de la similarité globale entre les régions se base sur la similarité des histogrammes des taux d'abondance. Cette étude permet d'aboutir à une décision concernant la fusion des régions « *Region Merging* » adjacentes ayant des caractéristiques des taux d'abondance similaires.

La Figure 6.23 présente une région R_1 et sept régions $R_2, R_3, R_4, R_5, R_6, R_7$ et R_8 adjacentes de cette région R_1 ainsi que les vecteurs descriptifs associés à chacune de ces régions. Notons qu'on a négligé les petites régions adjacentes de la région R_1 (par manque d'espace nécessaire pour présenter toutes les régions dans la même figure).

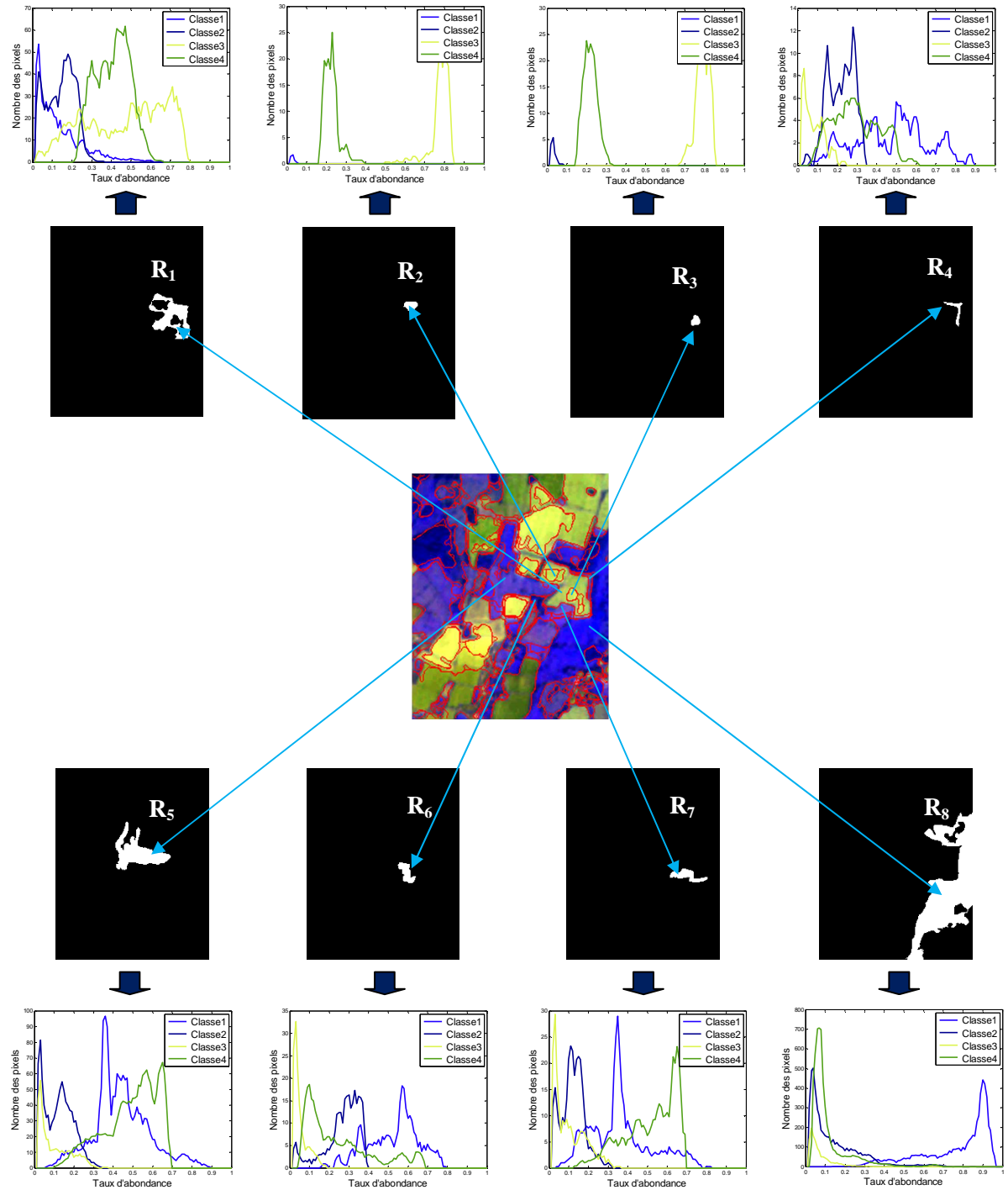


Figure 6.23 : Exemple de construction d'un rapport d'interprétation de scène

En considérant la fonction \mathcal{F} mesurant la similarité globale de régions, nous obtenons les résultats suivants :

- Pour les deux régions (R_1, R_2) : $\text{Sim}_G(V_{R_1}, V_{R_2}) = 0.13$
- Pour les deux régions (R_1, R_3) : $\text{Sim}_G(V_{R_1}, V_{R_3}) = 0.28$
- Pour les deux régions (R_1, R_4) : $\text{Sim}_G(V_{R_1}, V_{R_4}) = 0.26$
- Pour les deux régions (R_1, R_5) : $\text{Sim}_G(V_{R_1}, V_{R_5}) = 0.28$
- Pour les deux régions (R_1, R_6) : $\text{Sim}_G(V_{R_1}, V_{R_6}) = 0.12$

- Pour les deux régions (R_1, R_7) : $\text{Sim}_G(V_{R_1}, V_{R_7}) = 0.32$
- Pour les deux régions (R_1, R_8) : $\text{Sim}_G(V_{R_1}, V_{R_8}) = 0.14$

Pour la fusion des régions adjacentes, une contrainte est imposée concernant la similarité globale $\text{Sim}_G(V_{R_k}, V_{R_h})$ entre les régions adjacentes. Nous supposons que deux régions adjacentes peuvent être fusionnées si la similarité globale entre ces deux régions dépasse un seuil donné $\text{Sim}_G(V_{R_k}, V_{R_h}) \geq 0.3$ (valeur fixée d'une façon empirique). Par exemple, les deux régions R_1 et R_7 peuvent être fusionnées.

L'application des deux étapes précédentes (la fouille des régions et la fusion des régions) sur les huit régions $R_1, R_2, R_3, R_4, R_5, R_6, R_7$, et R_8 , nous permet de compléter le rapport d'interprétation par l'information suivante :

- La fusion des régions adjacentes : $\{(R_1, R_2) : \text{pas de fusion}, (R_1, R_3) : \text{fusion}, (R_1, R_4) : \text{pas de fusion}, (R_1, R_5) : \text{pas de fusion}, (R_1, R_6) : \text{pas de fusion}, (R_1, R_7) : \text{fusion}, (R_1, R_8) : \text{pas de fusion}\}$
- Taux d'abondance moyen des classes C_1, C_2, C_3 et C_4 dans les régions $R_1, R_2, R_3, R_4, R_5, R_6, R_7$, et R_8 :

	C_1		C_2		C_3		C_4	
	Centre d'inertie	Ecart type	Centre d'inertie	Ecart type	Centre d'inertie	Ecart type	Centre d'inertie	Ecart type
R_1	0.12	11.7	0.13	15.5	0.40	10.8	0.35	20.9
R_2	0.05	0.6	0	0	0.72	7.0	0.23	7.1
R_3	0	0	0.08	1.0	0.70	6.6	0.22	6.5
R_4	0.43	1.9	0.20	3.5	0.08	2.0	0.29	2.4
R_5	0.37	23.8	0.10	19.9	0.08	10.7	0.45	23.5
R_6	0.47	4.6	0.25	5.5	0.04	6.1	0.24	5.0
R_7	0.33	5.9	0.11	6.6	0.10	5.7	0.46	6.8
R_8	0.70	92.5	0.12	105.2	0.07	37.9	0.11	187.3

A partir de ce rapport, l'expert peut avoir une idée plus précise concernant la dispersion des classes dans les différentes régions identifiées :

- La classe C_1 a une forte présence dans la région R_8 et une présence moyenne dans les régions R_4, R_5, R_6 et R_7 .
- La classe C_2 a quasiment une faible présence dans les huit régions $R_1, R_2, R_3, R_4, R_5, R_6, R_7$, et R_8 .
- La classe C_3 a une forte présence dans les deux régions R_1 et R_2 et une présence moyenne dans la région R_3 et ceci malgré l'apparence visuelle de similarité entre ces trois régions.
- La classe C_4 a une présence moyenne dans toutes les régions sauf la région R_8 .

Nous constatons que la comparaison des grandes valeurs de taux d'abondance de classes dans les différentes régions permet de confirmer juste l'apparence visuelle de certaines de ces régions mais pas la similarité des histogrammes des taux d'abondance. Par exemple, la similarité globale de deux régions R_1 et R_7 , $\text{Sim}_G(V_{R_1}, V_{R_7}) = 0.32$ est supérieure de celle de deux régions R_1 et R_5 , $\text{Sim}_G(V_{R_1}, V_{R_5}) = 0.29$ malgré le fait que les grandes valeurs de taux d'abondance de classes dans les deux régions R_5 et R_7 soient très proches.

Il est aussi très important de noter les grandes valeurs de l'écart type de taux d'abondance des différentes classes dans certaines régions (R_1 , R_5 , et R_8) ce qui rend la classification des régions en fonction des forts taux d'abondance des classes très délicate. Par exemple, en se basant seulement sur cette valeur de taux d'abondance, la région R_1 sera attribuée à la classe C_3 malgré la présence très importante de la classe C_4 dans cette région.

6.5 Conclusion

Dans ce chapitre, nous avons présenté une extension de l'approche de segmentation/interprétation d'image proposée dans le cinquième chapitre. Cette extension a pour objectif de doter cette approche, d'un côté, d'opérations de fusion permettant d'exploiter plusieurs primitives dans la phase ascendante, et de l'autre côté, d'un système d'analyse de la scène réalisant principalement la fusion des régions détectées dans la phase ascendante.

Dans un premier temps, dans la phase ascendante, deux stratégies ont été considérées pour le processus de raffinement-estimation des distributions de possibilités : la première consiste à réaliser le raffinement-estimation image/image ou bien capteur/capteur et d'une façon indépendante et la deuxième qui consiste à réaliser le raffinement-estimation d'une façon conjointe (en prenant en compte toutes les images ou tous les capteurs). Nous avons montré que le raffinement-estimation conjoint apporte une homogénéité plus marquée des zones de forts degrés de possibilités pour les différentes classes.

L'élargissement réalisé dans la phase descendante de l'approche proposée a été fait par la construction des rapports descriptifs de chacune des régions identifiées dans la phase ascendante. Ces rapports comportent des informations caractéristiques sur chacune de ces régions (le vecteur descriptif, la localisation, la taille,...etc.). Le fait d'avoir ces rapports, nous a permis d'étudier deux types de la similarité entre les régions : la similarité liée à une classe thématique (qui permet d'explorer la dispersion de cette classe dans les différentes régions) et la similarité globale entre deux régions (qui permet de prendre une décision pour la fusion de deux régions adjacentes).

Nous avons évalué les performances de l'approche proposée sur une image multi-spectrale SPOT comportant trois bandes spectrales et quatre classes thématiques. Les résultats obtenus ont montré l'efficacité de l'approche proposée afin d'apprécier les régions en termes du contenu thématique et ceci malgré la différence de l'apparence visuelle entre ces régions. En conclusion, cette approche nous a permis de qualifier d'une façon plus approfondie le contenu thématique des différentes régions identifiées dans la scène.

CONCLUSIONS ET PERSPECTIVES

Dans ce travail de recherche, nous avons abordé la thématique de la représentation et du traitement possibiliste des connaissances en vue de concevoir un système d'interprétation de scène dans un environnement multi-sources de connaissances (i.e. capteurs, sources additionnelles de connaissances).

Le traitement d'informations imparfaites collectées à partir de la scène ainsi que la représentation des connaissances *a priori* sur cette scène, indispensable pour le processus d'interprétation, constituent des difficultés majeures rencontrées lors de la conception d'un système d'interprétation de scène. Afin de répondre à ses difficultés, nous avons choisi la théorie des possibilités, comme cadre de travail commun, pour la représentation et le traitement des informations et connaissances disponibles. En effet, la capacité de cette théorie à représenter des informations possibilistes (c'est-à-dire incertaines et décrites avec des connaissances épistémiques), ainsi que les mécanismes de raisonnement qu'elle propose pour permettre l'exploitation de telles informations et connaissances dites possibilistes, constituent les atouts essentiels justifiant le choix de cette théorie.

A. Interprétation de scène & théorie des possibilités

Dans le premier chapitre de ce manuscrit, le cadre général et les problématiques concernées par l'objectif de notre travail de recherche sur la conception d'un système d'interprétation de scène, ont été présentés.

Un bref rappel de certaines méthodes de reconnaissance de formes utilisées dans les systèmes d'interprétation de scène a été présenté. Ce rappel a montré l'intérêt des méthodes de reconnaissance de formes basées sur des modèles interprétables telles que les méthodes basées sur la théorie des ensembles flous ainsi que les méthodes basées sur la fusion de plusieurs sources de connaissances hétérogènes et ceci afin de rendre plus facile et être capable de résoudre les problèmes d'ambiguïté rencontrés lors d'un processus d'interprétation.

En fonction de l'objectif final visé par l'expert, un état de l'art autour des systèmes existants a été réalisé en les catégorisant selon cinq niveaux sémantiques d'interprétation (niveau sous-pixelique, niveau pixelique, niveau région, niveau objet et niveau scène globale). L'intérêt de cet état de l'art a été de montrer deux lacunes essentielles dans les systèmes d'interprétation de scène opérant sur les trois niveaux d'interprétation de scène les plus bas (sous-pixelique, pixelique et région). La première lacune réside dans le manque de liens explicites entre ces trois niveaux (c'est-à-dire l'interprétation des différentes régions homogènes identifiées au niveau région). La deuxième lacune est que la majorité des systèmes d'interprétation proposés reposent sur des connaissances *a priori* (un ensemble réduit de connaissances) pour la caractérisation des classes identifiées dans la scène d'où le besoin, lors du processus d'interprétation, d'acquisition des nouvelles connaissances à partir des connaissances disponibles avec une évaluation de la certitude des connaissances extraites

La caractérisation de l'information et les différentes formes d'imperfections pouvant l'entacher ainsi que les différents cadres mathématiques utilisés pour la représentation et le traitement de ces informations, ont été présentées dans le deuxième chapitre. Les quatre principales théories abordées sont:

la théorie des probabilités, la théorie des fonctions de croyance, la théorie des ensembles flous et la théorie des possibilités. Concernant cette dernière, différents concepts et outils tels que la distribution de possibilités, les mesures de possibilité et de nécessité, les règles de décision possibiliste, les opérateurs de fusion d'informations possibilistes ont été présentés. Ce chapitre a ainsi montré l'intérêt de la théorie des possibilités comme outil puissant permettant de gérer l'incertitude dans un contexte où les connaissances disponibles sont exprimées d'une façon ambiguë (les experts décrivent leurs connaissances par le biais de variables linguistiques, ou bien ces connaissances proviennent de l'utilisation d'un faible nombre d'échantillons extraits de la scène). Il est clair que tout processus d'interprétation de scène peut tirer profit des avantages procurés par une telle théorie, ceci la rend très utile dans un processus d'interprétation de scène.

B. Représentation possibiliste des connaissances extraites à partir des connaissances disponibles

Le troisième chapitre a réalisé une synthèse des différentes techniques d'estimation des distributions de possibilités, dans le domaine de l'interprétation de scène à partir d'images. En effet, ces distributions de possibilités encodent les connaissances liées à la description pixelique de classes thématiques présentes dans une image. Ces techniques d'estimation se répartissent en trois catégories selon qu'elles sont directement et explicitement guidées par l'expert, ou bien qu'elles reposent sur des processus d'apprentissage ou encore qu'elles transforment des distributions de probabilités en distributions de possibilités. Un intérêt particulier a été porté aux méthodes de transformation d'une distribution de probabilités en une distribution de possibilités car elle permet l'intégration des données statistiques. Parmi ces méthodes de transformation de données statistiques partielles (distribution de probabilités) en distribution de possibilités, la méthode proposée par Dubois-Prade permettant la transformation symétrique a été particulièrement détaillée et adoptée dans notre travail de recherche à cause des bonnes performances de cette méthode de transformation dans le domaine de la reconnaissance de formes.

L'objectif principal de ce chapitre a été de montrer la difficulté de caractériser les classes thématiques identifiées dans la scène à partir de variables linguistiques (ou de la sélection de formes standards afin d'estimer des distributions de possibilités) surtout dans un contexte de confusion interclasses où ces classes sont difficiles à discriminer (faible contraste, contexte très bruité, chevauchement spectral et spatial ou encore existence de classes multimodales).

Face à cette difficulté, l'expert tend à caractériser les classes via la délimitation, sur une image, des zones limitées considérées comme représentatives de ces classes. Une telle caractérisation impose un autre problème résidant dans l'estimation des distributions de possibilités à partir d'un ensemble limité des données statistiques extraites à partir de ces zones représentatives

Afin de répondre à ce problème de constitution de distributions de possibilités à partir d'un nombre réduit d'échantillons, une approche de raffinement des distributions de possibilités a été proposée dans ce chapitre. Cette approche est composée de trois étapes : une estimation des distributions de possibilités à partir d'un ensemble limité de données, une projection de ces distributions de possibilités sur l'image étudiée, et enfin, l'application d'un mécanisme de raisonnement possibiliste exploitant le contexte spatial comme étant une source additionnelle des connaissances permettant de sélectionner des nouveaux germes possibilistes appartenant avec une grande certitude aux différentes classes considérées dans la scène (qui seront intégrés dans l'ensemble des échantillons d'apprentissage d'une façon itérative). L'estimation des distributions de possibilités au cours du processus de raffinement a été réalisée en adoptant l'approche suivante : la technique KDE a été, d'abord, appliquée sur l'ensemble d'échantillons disponible afin d'estimer une distribution de probabilités et ensuite l'application de la transformation de

Dubois Prade symétrique a permis d'obtenir les distributions de possibilités pour les différentes classes thématiques présentes dans la scène analysée.

L'application de cette approche proposée (dans les deux contextes de faible et de grande complexité de discrimination entre les classes thématiques présentes dans la scène analysée) a montré que cette approche se focalise sur l'extraction des germes les plus discriminants (les germes les plus fréquents) de chaque classe thématique, et néglige l'ensemble des échantillons se trouvant dans la queue des distributions réelles représentant ces mêmes classes.

C. Notion de similarité entre informations possibilistes

Le quatrième chapitre a abordé l'utilisation de la notion de la similarité entre deux informations possibilistes (i.e. deux distributions de possibilités). Les différentes catégories existantes de mesures de similarité ont été présentées. Face au problème de choix d'une mesure de similarité, une méthode de sélection d'une mesure de similarité, en se basant sur la capacité de discrimination, en termes de classification pixelique, entre les classes thématiques contenues dans la scène, a été proposée.

L'objectif principal de ce chapitre était de présenter la notion de la similarité possibiliste et montrer son intérêt dans le processus d'interprétation de scène surtout son importance pour formaliser la discrimination entre les distributions de possibilités et évaluer ainsi l'ambiguïté due au « chevauchement » entre les classes.

Une démarche méthodologique permettant de réaliser la classification pixelique mais aussi l'opération de démixage pixelique en exploitant le concept de la similarité possibiliste pour l'intégration d'informations contextuelles spatiales et des connaissances *a priori* a été proposée. Dans cette démarche, chaque pixel à classer a été caractérisé en prenant en compte son voisinage au travers de la constitution de distributions de possibilités locales. Les deux processus de classification et de démixage pixelique ont été réalisés en exploitant la similarité entre distributions de possibilités locales et distributions de possibilités des classes, et non pas entre les degrés de possibilité d'appartenance des paramètres mesurés ponctuellement, aux différentes distributions de possibilités des classes.

Les résultats obtenus par l'application de cette démarche sur des images de synthèses ainsi que sur des images médicales réelles (en comparaison avec une approche Markovienne) semblent prometteurs.

D. Approche de segmentation/interprétation possibiliste multi-niveaux

Le cinquième chapitre a été consacré à la mise en œuvre de l'approche possibiliste de segmentation/interprétation d'image opérant sur les trois niveaux d'interprétation les plus bas (sous-pixelique, pixelique et région). Le fait de travailler sur les 3 niveaux d'interprétation conduit à une approche en deux phases : une phase ascendante qui permet de construire des régions de pixels homogènes à partir de pixels bruts (niveau pixelique → niveau région) suivie par une phase descendante qui exploite l'homogénéité des régions constituées pour analyser plus finement le contenu thématique des pixels sous-jacents (niveau région → niveau sous-pixelique).

Les connaissances injectées par l'expert constituent le point de départ de la phase ascendante. Cette phase regroupe l'approche de raffinement des distributions de possibilités et la segmentation par diffusion croisée des connaissances possibilistes qui résultent du processus de raffinement des distributions de possibilités. Ces connaissances sont représentées sous la forme de cartes possibilistes d'appartenance aux différentes classes. A l'issue de la phase ascendante, la scène analysée est donc représentée sous la forme d'une carte de régions homogènes.

La phase descendante de l'approche proposée comporte l'approche d'analyse et de démixage pixelique possibiliste. Dans cette phase, chaque région homogène identifiée dans la phase ascendante, est représentée par un vecteur descriptif comportant les histogrammes des taux d'abondance aux différentes classes de la scène analysée. L'objectif visé de cette phase est d'explicitier les différentes classes formant cette région ainsi que leur taux d'abondance.

Les résultats de l'évaluation ont montré un taux de reconnaissance intéressant (proche de celui obtenu par l'approche de segmentation markovienne) ainsi que la convergence rapide de la démarche possibiliste de segmentation d'image et l'efficacité de la démarche de démixage pixelique possibiliste en termes de caractérisation de chaque région d'intérêt. Cette évaluation a été réalisée en utilisant une base d'images de synthèse, des images mammographiques utilisées pour la détection de tumeurs du sein, ainsi que des images SONAR. La diversité des images réelles manipulées en termes de modalité d'acquisition mais aussi de scène sous-jacente, montre clairement le potentiel très intéressant de l'approche proposée.

E. Vers une interprétation globale de scène

Le dernier chapitre a été dédié à une extension de l'approche de segmentation/interprétation d'image proposée et ceci au cours de la phase ascendante et descendante. Dans la phase ascendante, deux stratégies de raffinement-estimation des distributions de possibilités ont été adoptées : raffinement-estimation image/image (et d'une façon indépendante) et la deuxième qui consiste à réaliser le raffinement-estimation d'une façon conjointe (en prenant en compte toutes les images). Nous avons montré que le raffinement-estimation conjoint apporte une homogénéité plus marquée des zones de forts degrés de possibilités pour les différentes classes. L'extension réalisée dans la phase descendante était de construire des rapports descriptifs de chacune des régions identifiées dans la phase ascendante et qui comportent des informations caractéristiques sur chacune de ces régions (le vecteur descriptif, la localisation, la taille,...etc.). Cette extension dans la phase descendante nous a permis de quantifier deux types de la similarité entre les régions : la similarité liée à une classe thématique (qui permet d'explorer la dispersion de cette classe dans les différentes régions) et la similarité globale entre deux régions (qui permet de prendre une décision pour la fusion de deux régions adjacentes).

Une évaluation globale de l'approche proposée a été réalisée en utilisant une image multi-spectrale SPOT comportant trois bandes spectrales et quatre classes thématiques. Les résultats obtenus ont montré l'efficacité de l'approche proposée pour qualifier d'une façon plus approfondie le contenu thématique des différentes régions identifiées dans la scène.

F. Bilan

Ce travail de thèse propose une approche globale contribuant aux objectifs visés par tout problème *difficile* d'interprétation de scène. Cette approche opérant aux trois niveaux d'interprétation sous-pixelique, pixelique et région, et dans un environnement multi-sources de connaissances, a permis d'exploiter les complémentarités intrinsèques qui existent entre ces trois niveaux afin d'offrir des outils possibilistes permettant d'améliorer l'interprétation des données acquises et ainsi, d'aboutir à une représentation plus riche de la scène observée. Dans cette approche, la sélection des zones représentatives des classes thématiques présentes dans l'image constitue la seule intervention de l'expert.

Les résultats obtenus par l'application de l'approche proposée sur plusieurs types d'images (images de synthèse, images mammographiques, images du type IRM, images SONAR et images satellitaires) ont été très encourageants et ont montré l'efficacité de la théorie des possibilités comme un cadre de représentation des connaissances *a priori* de l'expert mais aussi de raisonnement et de fusion permettant

d'une part d'extraire de nouvelles connaissances à partir des connaissances disponibles, et d'autre part d'intégrer plusieurs sources de connaissances dans le processus d'interprétation de scène.

G. Perspectives

Notre travail de thèse a permis d'ouvrir plusieurs perspectives. Ces perspectives peuvent être regroupées en deux catégories :

A- Perspectives à court terme

- La mise en place d'une coopération entre l'approche de segmentation proposée dans la phase ascendante et d'autre méthodes, dites *contours*, peut être envisagée. Une telle coopération ne peut qu'améliorer les résultats globaux de la segmentation surtout dans les zones frontières. De même, dans les cas des images multi-spectrales et hyper-spectrales, une coopération peut être effectuée entre la méthodologie de démixage proposée dans la phase descendante et d'autres méthodes de démixage proposée dans la littérature.
- Le choix des valeurs de seuil possibiliste (pour la sélection des germes possibilistes) a été réalisé en utilisant des images de synthèse et d'une façon empirique. Ce choix mérite encore des recherches dans le but de développer une méthodologie permettant de choisir ce seuil d'une façon automatique et selon l'application envisagée (en se basant sur les zones d'apprentissage, par exemple).

B- Perspectives à long terme

- L'analyse hiérarchique reposant sur les deux phases ascendante et descendante et opérant aux niveaux d'interprétation les plus bas (sous-pixelique, pixelique et région) ne représente qu'une instance d'une approche générique opérant sur trois niveaux voisins d'interprétation. Ainsi, ce concept d'analyse hiérarchique pourrait être répliqué et adapté à des niveaux d'abstraction plus élevés (niveau objet et niveau scène). Par exemple, une phase ascendante partant du niveau objet vers le niveau scène, permet d'identifier les objets d'intérêt et éventuellement leurs relations spatiales avec d'autres objets dans la scène alors qu'une phase descendante partant de niveau scène vers le niveau région, permet d'identifier les différentes régions composant ces objets d'intérêt et ainsi la possibilité de reconnaître des sous-classes d'objets (Figure 7.1).
- Dans son état actuel, le système d'interprétation de scène proposé ne prend pas en compte une dimension fondamentale de l'interprétation de scène à savoir la dimension temporelle qui permet la compréhension de l'évolution de cette scène et au travers de cette évolution, une compréhension de la scène observée à un instant donné. L'extension envisagée serait de permettre au système proposé d'effectuer une interprétation de séquences d'images de la scène observée (i.e. en termes d'événements). En effet, une telle extension peut être réalisée par l'étude de l'évolution des vecteurs descriptifs des régions d'intérêt et qui peut permettre l'analyse, l'interprétation et le suivi plus spécifique de changement dans la scène. Néanmoins, l'introduction de la composante temporelle nécessite également la prise en compte de nouvelles problématiques telles que les différences de conditions de prises des images, etc.

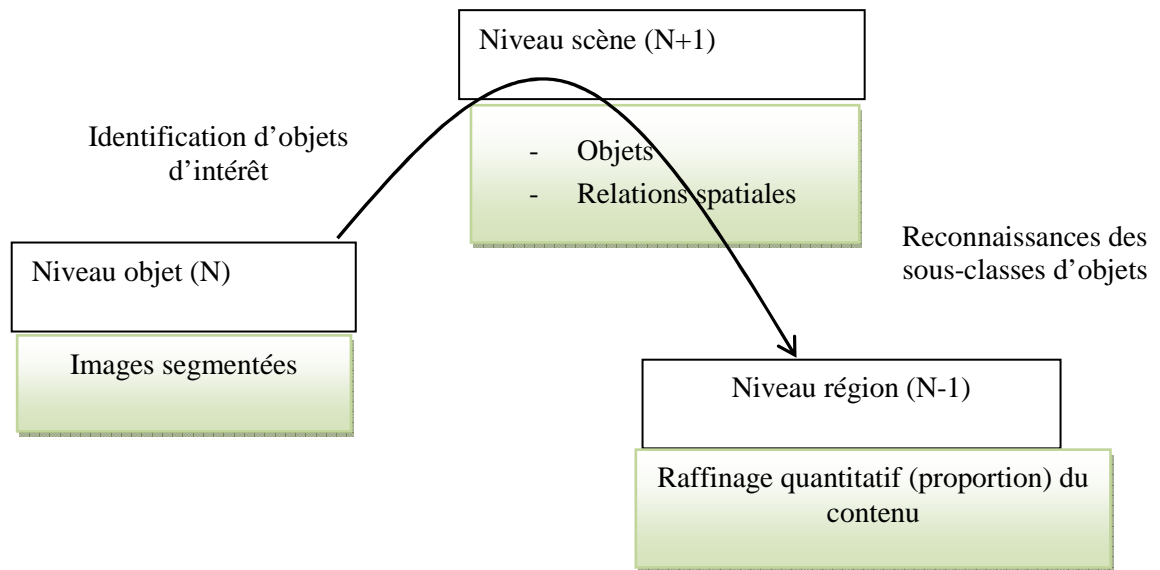


Figure 7.1 : Brique conceptuelle générique exploitant la richesse des connaissances portant sur des niveaux sémantiques voisins N-1, N et N+1 (avec N correspondant ici au niveau objet)

Références

- [1] C. Hudelot, "Towards a cognitive vision platform for semantic image interpretation, application to the recognition of biological organisms", thèse, *université de Nice - Sophia Antipolis, France*, 2005.
- [2] D. A Forsyth and J. Ponce, "Computer vision, a modern approach", 2nd ed, *Prentice Hall*, 2003.
- [3] Z. B. Rabah, "Demixage spectral d'images hyperspectrales et fusion possibiliste des connaissances : application a l'érosion hydrique", thèse, *telecom Bretagne, France*, 2011.
- [4] D. Marr, "VISION". *W. H. Freeman and Company, New York*, 1982.
- [5] A. Boucher and T. Lee, "Comment extraire la sémantique d'une image ? ", *the 3rd International Conference : Sciences of Electronic, Technologies of Information and Telecommunication*, Tunisia, 2005.
- [6] A.W.M. Smeulders, M. Worring, S.Santini, A.Gupta, and R. Jain, "Content-based image retrieval at the end of the early years", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp.1349-1380, 2000.
- [7] P. Garnesson, G. Giraudon, and P. Montesinos, "MESSIE: un système multi spécialistes en vision. Application à l'interprétation en imagerie aérienne", *Traitement de signal*, vol. 9, pp. 403-419,1992.
- [8] T. Matsuyama, "Expert systems for image processing: knowledge-based composition of image analysis processes", *Computer Vision, Graphics and Image Processing*, vol. 48, pp.22-49, 1980.
- [9] I. Bloch and A. Hunter, "Fusion: General concepts and characteristics", *International Journal of Intelligent System*, vol. 16, pp. 1107-1134, 2001.
- [10] V. Clement, G. Giraudon, S. Houzelle, and F. Sandakly, "Interpretation of remotely sensed images in a context of multisensor fusion using a multispecialist architecture", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 779-791, 1993.
- [11] I. Bloch, "Information combination operators for data fusion: A comparative review with classification," *IEEE Transaction on Systems, Man, and Cybernetic*, vol. 26, pp. 52-67, 1996.
- [12] D. Lu and Q. Wang, "A survey of image classification methods and techniques for improving classification performance", *Int. Journal of Remote Sensing*, vol. 28, pp. 823-870, 2007.
- [13] N. Keshava, "A survey of spectral unmixing algorithms", *Lincoln laboratory journal*, vol. 14, pp.55-78, 2003.
- [14] I. Bloch, "Fuzzy spatial relationships for image processing and interpretation: a review", *Image and Vision Computing*, vol. 23, pp.89-110, 2005.
- [15] V. Amberg, "Analyse de scènes péri-urbaines à partir d'images radar haute résolution. Application à l'extraction semi-automatique du réseau routier", thèse, *Institut de Recherche en Informatique de Toulouse, France*, 2005.
- [16] A. R. Rao and R. Jain, "Knowledge representation and control in computer vision systems", *IEEE Expert*, vol. 3, pp. 64-79, 1988.
- [17] T. Matsuyama and V. Hwang, "SIGMA: A framework for image understanding, integration of bottom-up and top-down analyses", *Proc. of 9th IJCAI*, pp.908-915, 1985.

- [18] S. Magnussen, P. Boudewyn, and M. Wulder, "Contextual classification of Landsat TM images to forest inventory cover types", *International journal of remote sensing*, vol. 25, pp. 2421-2440, 2004.
- [19] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp.721-741, 1984.
- [20] B. Tso and R. C. Olsen, "A contextual classification scheme based on MRF model with improved parameter estimation and multiscale fuzzy line process", *Remote Sensing of Environment*, vol. 97, pp. 127-136, 2005.
- [21] J. Besag, "On the statistical analysis of dirty pictures", *Journal of the Royal Statistical Society*, vol. 48, pp. 259-302, 1986.
- [22] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian Mixture Models to Structured Sparsity", *IEEE Transactions on Image Processing*, vol. 21, pp. 2481-2499, 2012.
- [23] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society*, vol. 39, pp. 1-38, 1977.
- [24] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors", *Nature*, vol. 323, pp. 533-536, 1986.
- [25] G. Cybenko, "Approximation by superpositions of a sigmoidal function", *Mathematics of Control, Signals, and Systems*, vol. 2, pp. 303-314, 1989.
- [26] J. M. Keller, "Fuzzy Logic in Computer Vision", *Proceeding of the 6th International Fuzzy System Association – IFSA95*, pp. 7-10, 1996.
- [27] R. M. Haralick, M. Shanmugam, and I. Dinstein, "Texture feature for image classification", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, pp. 610-621, 1973.
- [28] L. Wald, "Some terms of reference in data fusion", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 1190-1193, 1999.
- [29] I. R. Farah, W. Boulila, K. S. Ettabaa, B. Solaiman, and M. Ben Ahmed, "Interpretation of Multisensor Remote Sensing Images: Multiapproach Fusion of Uncertain Information", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 4142-4152, 2008.
- [30] D. Crevier and R. Lepage, "Knowledge-based image understanding systems: a survey", *Computer. Visison.and Image Understing*, vol. 67, pp.160-185, 1997.
- [31] R. D. Rimey and C. M. Brown, "Control of selective perception using bayes nets and decision theory", *International Journal of Computer Vision*, vol. 12, pp.173-207, 1994.
- [32] P. Perner, "Why case-based reasoning is attractive for image interpretation", *In ICCBR*, pp. 27-44, 2001.
- [33] L. Bruzzone and D. F. Prieto, "A Bayesian approach to automatic change detection", *International Geoscience and Remote Sensing Symposium*, vol. 3, pp. 1816-1818, 1999.
- [34] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Hyperspectral image classification with independent component discriminant analysis", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 4865-4876, 2011.
- [35] R. Duda, P. Hart, and D. Stork, "Pattern Classification", 2nd ed, New York: Wiley, 2001.
- [36] F. Ludic, "Comparison of several fusion paradigms applied to pixel-based image classification", *International Conference on Information Fusion*, pp. 1-8, 2006.
- [37] L. Roux and J. Desachy, "Satellite image classification based on multi-Source information-fusion with possibility theory", *International Geoscience and Remote Sensing Symposium*, vol. 2, pp. 885-887, 1994.
- [38] B. Solaiman, L. E. Pierce and F. T. Ulaby, " Multisensor data fusion using fuzzy concepts: application to land-cover classification using ERS-1/JERS-1 SAR composites", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 1316-1326, 1999.
- [39] X. Jia and L. Wang, "Fuzzy assessment of spectral unmixing algorithms", *IEEE Journal of Selected*

- Topics in Applied Earth Observations and Remote Sensing*, pp.1-9, 2013.
- [40] S. Verbeiren, H. Eerens, and J. V. Orshoven, "Sub-pixel classification of spot-vegetation time series for the assessment of regional crop areas in Belgium", *International Journal of Applied Earth Observation and Geoinformation*, vol. 10, pp.486-497, 2008.
 - [41] F. Liu, D. Xu, C. Yuan, and W. Kerwin "Image segmentation based on Bayesian network-Markov Random Field and its application to in vivo plaque composition", *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro*, pp. 141-144, 2006.
 - [42] R. Azmi and N. Norozi "A new markov random field segmentation method for breast lesion segmentation in MR images", *Journal of Medical Signals and Sensors*, vol. 1, pp. 156-164, 2011.
 - [43] F. Prandi, R. Brumana and F. Fassi, "Semi-automatic objects recognition in urban areas based on fuzzy logic", *Journal of Geographic Information System*, vol. 2, pp. 55-62, 2010.
 - [44] B. Neumann and T. Weiss, "Navigating through logic-based scene models for high-level scene interpretations", *International Conference on Computer Vision Systems - ICVS*, pp. 212-222, 2003.
 - [45] B. Gong, J. Liu, X. Wang, and X. Tang, "Learning semantic signatures for 3D object retrieval", *IEEE Transactions on Multimedia*, vol. 15, pp. 369-377, 2013.
 - [46] P. Sobrevilla and E. Montseny, "Fuzzy sets in computer vision: an overview", *Mathware and Soft Computing*, vol. 10, pp.71-83, 2003.
 - [47] C. Brewster, K. O'Hara, S. Fuller, Y. Wilks, E. Franconi, M.A. Musen, J. Ellman and S.B. Shum, "Knowledge Representation with Ontologies: The Present and Future", *IEEE Intelligent Systems*, vol. 19, pp. 72-81, 2004.
 - [48] E. C. C.Tsang, D. S. Yeung, and J. W. T. Lee, "Refinement of knowledge representation parameters in fuzzy production rules by genetic algorithms", *IEEE International Conference on Systems Man Cybernetics*, vol. 2, pp. 1518-1523, 1998.
 - [49] S. Peters and H. Shrobe, "Using semantic networks for knowledge representation in an intelligent environment", *IEEE International Conference on Pervasive Computing and Communications*, pp. 323-329, 2003.
 - [50] C. Zins, "Conceptual approaches for defining data, information, and knowledge", *Journal of the American Society for Information Science and Technology*, vol. 58, pp. 479-493, 2007.
 - [51] B. Solaiman, "Information fusion concepts: from information elements definition to the application of Fusion Approaches", *SPIE proceedings series*, vol. 4385, pp. 205-212, 2001.
 - [52] L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets Systems*, vol. 1, pp. 3-28, 1978.
 - [53] I. Bloch and H. Maître "Fusion en traitement d'images: spécificités et approches numériques", *Traitement du signal : bases théoriques, TE5230*, pp. 1-23, 2002.
 - [54] I. Bloch "Fusion d'informations en traitement du signal et des images", *Hermes Science Publication*, Janvier 2003.
 - [55] A. N. Kolmogorov, "foundation of the theory of probability", 2nd ed. New York: Chelsea Publishing Company, 1956.
 - [56] R. O. Duda and P. E. Hart, "Pattern Classification and Scene Analysis", John Wiley & Sons Inc, 1973.
 - [57] P. Walley, "Measures of uncertainty in expert systems", *Artificial Intelligence*, vol. 83, pp. 1-58, may 1996.
 - [58] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping". *Annals of Mathematical Statistics*, vol. 38, pp. 325-339, 1967.
 - [59] G. Shafer, "Mathematical theory of evidence". Princeton Univ Pr, 1976.
 - [60] L. Zadeh, "Fuzzy sets", *Information and Control*, vol. 8, pp. 338-353, 1965.
 - [61] M. Ji, "Using fuzzy sets to improve cluster labeling in unsupervised classification", *International Journal of Remote Sensing*, vol. 24, pp. 657-671, 2003.

- [62] S. N. Mandal, J. P. Choudhury, D. De, and S. R. B. Chaudhuri, "Roll of membership functions in fuzzy logic for prediction of shoot length of Mustard plant based on residual analysis", *World Academy of Science, Engineering and Technology*, vol. 40, pp. 378-384, 2008.
- [63] B. Tso and P. M. Mather, "Classification methods for remotely sensed data", 2nd ed, *Taylor & francis group*, 2009.
- [64] J. A. Benediktsson, P.H. Swain, and O.k. Esroy, "Neural network approaches versus statistical methods in classification of multisource remote sensing data", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 540-552, 1990.
- [65] G. M. Foody, R. M. Lucas, P. J. Curran, and M. Honzak, "Non-linear mixture modelling without end-members using an artificial neural network", *International Journal of Remote Sensing*, vol. 18, pp. 937-953, 1997.
- [66] E. C. B. De Colstoum and C. L. Walthall, "Improving global scale land cover classifications with multi-directional POLDER data and a decision tree classifier", *Remote Sensing of Environment*, vol. 100, pp. 474-485, 2006.
- [67] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition", *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, 1998.
- [68] O. Demirkaya, M. H. Asyali, and P.K. Sahoo, "Imag processing with matlab, application in medicine and biology", *Taylor & francis group*, 2009.
- [69] D. Dubois and H. Prade, "Possibility theory: an approach to computerized processing of uncertainty", *Kluwer Academic / Plenum Publishers*, 1988.
- [70] D. Dubois and H. Prade, "Fuzzy Sets and Systems: Theory and Applications", *Academic Press*, New York, 1980.
- [71] L. A. Klein, "Sensor and data fusion: a tool for information assessment and decision making", 2nd ed, *SPIE-the international society for optical engineering*, Washington, 2007.
- [72] L. Roux, "An application of possibility theory information fusion to satellite image classification", *Fuzzy Logic in Artificial Intelligence*, pp. 166-179, 1999.
- [73] D. Dubois and H. Prade, "La fusion d'informations imprécises", *Traitement du signal*, vol. 11, pp. 447-458, 1994.
- [74] M. H. Alsun, "Indexation guidée par les connaissances en imagerie medicale", thèse, *telecom Bretagne, France*, 2012.
- [75] I. Bloch: "Information combination operators for data fusion: a comparative review with classification", *IEEE. Transactions on Systems, Man, and Cybernetics*, vol. 26, pp. 52-67, 1996.
- [76] Sh. Kikuchi and V. Perincherry, "Handling Uncertainty in Large Scale Systems with Certainty and Integrity," *MIT Engineering Systems Symposium*, Cambridge, 2004.
- [77] S. Medasani, J. Kim, and R. Krishnapuram, "An overview of membership function generation technique for pattern recognition", *International Journal of Approximate Reasoning*, vol. 19, pp. 391-417, 1998.
- [78] D. Dubois and H. Prade, "Fuzzy sets - a convenient fiction for modeling vagueness and possibility", *IEEE Transactions on Fuzzy Systems*, vol. 2, pp.16-21, 1994.
- [79] R. Krishnapuram, J.M. Keller, and Y. Ma, "Quantitative analysis of properties and spatial relations of fuzzy image regions", *IEEE Transactions on Fuzzy Systems*, vol. 1, pp. 222-233, 1993.
- [80] I. Bloch, "Fuzzy relative position between objects in image processing: a morphological approach", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 657-664, 1999.
- [81] C. R. Mihalache and M. Craus, "Neural network and fuzzy membership functions based edge detection for digital images", *16th International Conference on System Theory, Control and Computing (ICSTCC)*, pp. 1-6, 2012.
- [82] S. Auenphanwiriyaikul, S. Attrapadung, S. Thovuttikul and N. Theera-Umpon, "Breast abnormality detection in mammogram using fuzzy inference system", *14th International Conference on Fuzzy System*,

- pp. 155-160, 2005.
- [83] T. Nakashima, G. Schaefer, Y. Yokota, and H. Ishibuchi, "A weighted fuzzy classifier and its application to image processing tasks", *Fuzzy Sets and Systems*, vol. 158, pp. 284-294, 2007.
 - [84] H. D. Cheng and J.R. Chen, "Automatically determine the membership function based on the maximum entropy principle", *Information Sciences*, vol. 96, pp. 163-182, 1997.
 - [85] G. Nieradka and B. Butkiewicz, "A method for automatic membership function: estimation based on fuzzy measures", *International Fuzzy Systems Association (IFSA)*, pp. 451-460, 2007.
 - [86] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", *Science*, vol. 220, pp. 671 -680, 1983.
 - [87] D. E. Goldberg, "Genetic algorithm in searching, optimization, and machine learning", *Addison-Wesley, Reading*, 1989.
 - [88] J. Kennedy and R. Eberhart, "Particle swarm optimization", *Proceedings of IEEE International Conference on Neural Networks*, pp. 1942-1948, 1995.
 - [89] M. Zhao, "Image thresholding techniques based on fuzzy partition and entropy maximization", *University of Sydney. School of Electrical and Information Engineering*, 2006.
 - [90] D. Dubois and H. Prade, "Unfair coins and necessity measures: towards a possibilistic interpretation of histograms", *Fuzzy Sets and Systems*, vol. 10, pp. 15-20, 1983.
 - [91] D. Dubois, L. Foulloy, G. Mauris, and H. Prade, "Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities", *Reliable computing*, pp. 273-297, 2004.
 - [92] G. Klir, "A principle of uncertainty and information invariance", *International Journal of General Systems*, vol. 17, pp. 249-275, 1990.
 - [93] K. Yamada, "Probability-possibility transformation based on evidence theory", *IFSA World Cong. and 20th NAFIPS Inter. Conf.*, vol. 1, pp. 70-75, 2001.
 - [94] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms", *Plenum Press, New York*, 1981.
 - [95] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition", *Medical Physics*, vol. 20, pp. 1033-1048, 1993.
 - [96] R. Krishnapuram and J. M. Keller, "A possibilistic approach for clustering", *IEEE Transactions on Fuzzy Systems*, vol. 1, pp.98-110, 1993.
 - [97] N. Ferahta, A. Moussaoui, K. Benmahammed, and V. Chen, "New fuzzy clustering algorithm applied to RMN image segmentation", *International Journal of Soft Computing*, vol.1, pp. 137-142, 2006.
 - [98] R. O. Duda and P. E. Hart, "Pattern Classification and Scene Analysis". *New York: Wiley*, 1973.
 - [99] J.M. Keller, M.R. Gray, and J.A. Givens, "A fuzzy K-nearest neighbor algorithm", *IEEE Transactions on Systems, Man, and Cybernetics*, vol.15, pp. 580-585, 1985.
 - [100] F. Huet and S. Philipp, "Fusion of images interpreted by a new fuzzy classifier", *Pattern Analysis & Applications*, vol 1, pp. 231-247, 1998.
 - [101] B. Charroux, "Analyse d'images: coopération d'opérateurs de segmentation guidée par l'interprétation", *PhD thesis, University Paris XI Orsay, France*, 1996.
 - [102] J.Y. Chang and C. W. Cho, "Scene analysis system using a combined fuzzy logic-based technique", *Journal of the Chinese Institute of Engineers*, vol. 25, pp. 297-307, 2002.
 - [103] E. Hüllermeier, "Possibilistic Instance-Based Learning", *Artificial Intelligence: Fuzzy set and possibility theory-based methods in artificial intelligence*, vol. 148, pp. 335-383, 2003.
 - [104] B. Alsahwa, S. Almouahed, D. Gueriot, S. Daniel and B. Solaiman, "classification d'images basée sur l'utilisation de la similarité possibiliste", *TAIMA11, 03-08 octobre, Hammamet, Tunisie*, 2011
 - [105] B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, "Possibilistic similarity based Image

- classification”, *ICPRAM2013, 15-18 February, Barcelone, Spain*, 2013.
- [106] B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “Iterative possibility distributions refining in pixel-based images classification framework”, *ICPRAM2013, 15-18 February, Barcelone, Spain*, 2013
 - [107] E. Parzen, “On estimation of probability density function and mode”. *Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, 1962.
 - [108] B. W. Silverman, “Density Estimation for Statistics and Data Analysis”, *London: Chapman & Hall/CRC*, 1998.
 - [109] L. Bentabet, S. Jodouin, and A. Boudraa, “Estimation of mass functions in Dempster–Shafer evidence theory using fuzzy clustering and spatial information for gray level based image fusion,” *Optical Engineering*, vol. 41, pp. 760-770, 2002.
 - [110] Sayed Mouchaweh, “Semi-supervised classification method for dynamic applications”, *Fuzzy Sets and Systems*, vol. 161, pp. 544-563, 2010.
 - [111] W. Eziddin, “Segmentation itérative d’images par propagation de connaissances dans le domaine possibiliste”, *Thèse, Télécom Bretagne, France*, Juin 2012.
 - [112] B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “Contextual possibilistic knowledge diffusion for images classification” *Image processing and Communications Challenges 5*, vol. 233, pp. 161-169, 2014.
 - [113] R. B. Dubey, M. Hanmandlu and S. K. Gupta, “Level set detected masses in digital mammograms”, *Indian Journal of Science and Technology*, vol. 3, pp. 9-13, 2010.
 - [114] H. Zhang, S. W. Foo, S. M. Krishnan and C. H. Thng, “Automated breast masses segmentation in digitized mammograms”, *IEEE International Workshop on Biomedical Circuits and Systems*, pp. S2/2- S1-4, 2004.
 - [115] R. Adams and L. Bischof, “Seeded region growing”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.16, pp. 641-647, 1994.
 - [116] T.Gadi and R. Benslimane, “Segmentation hiérarchique floue”, *Traitement du signal*, vol. 17, pp. 59-67, 2000.
 - [117] M. Nagao and T. Matsuyama, “Edge Preserving Smoothing”, *Computer Graphics and Image Processing* 9, pp. 394-407, 1979.
 - [118] J. Durkin, “expert systems design and development”, *Prentice Hall*, 1994.
 - [119] I. Jurisica and J. Glasgow, “Applications of case-based reasoning in molecular biology”, *AI Magazine*, vol. 25, pp. 85-95, 2004.
 - [120] D. Hume, “An enquiry concerning human understanding”, *Oxford University Press Inc., New York*, 1999.
 - [121] B. Bouchon-Meunier, M. Rifqi, and M-J. Lesot, “Similarities in fuzzy data mining: from a cognitive view to real-world applications”, *Springer LNCS*, vol. 5050, pp. 349-367, 2008.
 - [122] E. L. Rissland, “AI and Similarity”, *IEEE Intelligent Systems*, vol. 21, pp. 39-49, 2006.
 - [123] J-P.Vert, K. Tsuda, B. Schölkopf , “A primer on kernel methods”, *Kernel Methods in Computational Biology*, 2004.
 - [124] G. Gan, C. Ma, et J. Wu, “Data clustering: theory, algorithms, and applications”, *Society for Industrial and Applied Mathematics*, 2007
 - [125] R.C. Amorim and B. Mirkin, “Minkowski metric, feature weighting and anomalous cluster initialisation in K-Means Clustering”, *Pattern Recognition*, vol. 45, pp. 1061-1075, 2012.
 - [126] A. Zemirline, “Définition et fusion de systèmes diagnostic l’aide d’un processus de fouille de données: application aux systèmes diagnostics”, thèse, *University of Rennes 1*, 2008.
 - [127] G. Gan, C. Ma, and J. Wu, “Data clustering: theory, algorithms, and applications”, *Society for Industrial and Applied Mathematics*, pp. 25-41, 2007.
 - [128] W. Pedrycz, “Knowledge-based clustering: from data to information granules”, *John Wiley and Sons*, pp.

- 1-27, 2005.
- [129] I. Jenhani, "From possibilistic similarity measures to possibilistic decision trees", thèse, *Université de Tunis, Université d'Artois*, 2010.
 - [130] G. Bisson, "La similarité: une notion symbolique/ numérique", *Apprentissage Symbolique/ Numerique*, (tome 2), Eds. Moulet Brito, Edition CEPADUES, pp. 169-201, 2000.
 - [131] A. Tversky, "Features of similarity", *Psychological Reviews*, vol. 84, pp. 327-352, 1977.
 - [132] R. Zwick, E. Carlstein, and D.V. Budeanu, "Measures of similarity among fuzzy concepts: A comparative analysis", *International Journal of Approximate Reasoning*, vol. 1, pp. 221-242, 1987.
 - [133] Y. Li, J.-M. Liu, J. Li, W. Deng, C.-X. Ye, and Z.-F. Wu "The fuzzy similarity measures for content-based image retrieval" *International Conference on Machine Learning and Cybernetics*, vol. 5, pp. 3224-3228, 2003.
 - [134] R. Jain, S.N.J. Murthy, P.L.-J. Chen, and S. Chatterjee, "Similarity measures for image databases", *IEEE International Conference on Fuzzy Systems*, pp. 1247-1254, 1995.
 - [135] B. Bouchon-Meunier, M. Rifqi, and S. Bothorel, "Towards general measures of comparison of objects", *Fuzzy Sets and Systems*, vol 84, pp.143-153, 1996.
 - [136] B.Bouchon-Meunier, G. Coletti, M. Lesot, and M. Rifqi, "Towards a conscious choice of a similarity measure: A qualitative point of view", *The 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pp. 542-553, 2009.
 - [137] C. Mencar , G. Castellano , A. M. Fanelli , and A. Bargiela, "Similarity vs. possibility in measuring fuzzy sets distinguishability", *International Conference on Recent Advances in Soft Computing*, Nottingham, UK, 2004.
 - [138] A. S. Dewalle-Vignion, N. Betrouni, R. Lopes, D. Huglo, S. Stute, and M. Vermandel, "A new method for volume segmentation of PET images, based on possibility theory", *IEEE Transactions on Medical Imaging*, vol. 30, pp. 409-423, 2011.
 - [139] L. Chaabane and M. Abdelouahab, "Segmentation of MR Brain Images using a Data Fusion Approach", *International Journal of Computer Applications*, Volume 36, pp. 27-32, 2011.
 - [140] L. Wang, H. Ji, and X. Gao "MR brain image segmentation using a possibilistic entropy based clustering method" *International Conference on Signal Processing*, vol.3, pp. 2241-2244, 2004.
 - [141] I. Jenhani, N. Ben Amor, Z. Elouedi, S. Benferhat, K. Mellouli, "Information affinity: A new similarity measure for possibilistic uncertain information", *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pp. 840-852, 2007.
 - [142] C-I Chang and A. Plaza, "Fast iterative algorithm for implementation of pixel purity index", *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 63-67, 2006.
 - [143] F. Kruse, A. Lefkoff, J. Boardman, K. Heiderbrecht, P.J. Shapiro, and A.F.H. Goetz, "The spectral image processing system - interactive visualisation and analysis of imaging spectrometer data", *Remote Sensing of Environment*, vol. 44, pp. 145-163, 1993.
 - [144] W. Boardman, "Inversion of imaging spectrometry data using singular value decomposition", *Geoscience and Remote Sensing Symposium*, pp. 2069-2072, 1989.
 - [145] J.J. Settle, "On the relationship between spectral unmixing and subspace projection", *IEEE Transaction on Geoscience and Remote Sensing*, vol. 34, pp. 1045-1046, 1996.
 - [146] J. Herault and C. Jutten, "Blind separation of sources, part I : An adaptive algorithm based on neuromimetic architecture", *Signal Processing*, vol. 24, pp.1-10, 1991.
 - [147] B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, "A method of pixel unmixing by classes based on possibilistic similarity", *ICPRAM2014*, 6-8 March, Angers, France, 2014
 - [148] Z. Liu, J. Dezert, G. Mercier, and Q. Pan, "Dynamic evidential reasoning for change detection in remote sensing images", *IEEE Transactions on geoscience and remote sensing*, vol. 50, pp. 1955-1967, 2012.
 - [149] D. Van der Weken, M. Nachtegaal, and E. Kerre, "Some new Similarity measures for histograms",

- ICVGIP 2004*, 16-18 December, Kolkata, India, 2004.
- [150] F. D. Jou, K. C. Fan, and Y.L. Chang, "Efficient matching of large-size histograms", *Pattern Recognition Letters*, vol. 25, pp. 277-286, 2004.
- [151] M. Swain and D. Ballard, "Color indexing", *International Journal of Computer Vision*, vol. 7, pp. 11-32, 1991.

Publications

Conférences

1. B. Alsahwa, S. Almouahed, D. Gueriot, S. Daniel and B. Solaiman, “Classification d’images basée sur l’utilisation de la similarité possibiliste”, *TAIMA11*, 03-08 octobre, Hammamet, Tunisie, 2011.
2. B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “Possibilistic similarity based Image classification”, *ICPRAM2013*, 15-18 february, Barcelone, Spain, 2013.
3. B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “Iterative possibility distributions refining in pixel-based images classification framework”, *ICPRAM2013*, 15-18 february, Barcelone, Spain, 2013.
4. B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “Contextual possibilistic knowledge diffusion for images classification” *Image processing and Communications Challenges 5*, vol. 233, pp. 161-169, 2014.
5. B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “A method of pixel unmixing by classes based on the possibilistic similarity”, *ICPRAM2014*, 5-8 March, Angers, France, 2014.

Revues en cours

1. B. Alsahwa, S. Almouahed, D. Gueriot, and B. Solaiman, “Possibilistic knowledge refining and diffusion for images segmentation”, *Fuzzy sets and systems*.
2. B. Solaiman, S. Almouahed, B. Alsahwa, and D. Gueriot, “Possibilistic decision making in pattern recognition systems”, *IEEE Transactions on Fuzzy Systems*.

Résumé

Cette thèse s'inscrit dans le contexte de la représentation et du traitement possibiliste des connaissances en vue de concevoir un système d'interprétation de scène observée par de multiples capteurs. Les capteurs fournissant une ou des informations imparfaites, les recherches se sont focalisées sur une représentation possibiliste des connaissances. En effet, ce formalisme permet d'introduire et de manipuler explicitement au sein d'un même cadre, certaines imperfections portées par les informations collectées mais aussi certaines connaissances a priori sur la scène, les capteurs, les conditions d'acquisition exprimées de manière vague soit par des experts soit résultant d'observations ou de jugements ambigus.

Ce travail est divisé en plusieurs étapes : état de l'art et maîtrise de la théorie des possibilités, analyse hiérarchique/sémantique de scène, étude de l'exploitation de la théorie des possibilités dans différents niveaux sémantiques (sous-pixelique, pixelique, objet et scène) et finalement, application de l'approche proposée à des fins d'interprétation de scène.

L'objectif final de cette thèse est de développer un système d'interprétation de scène basé sur une démarche possibiliste. Cette approche permet d'effectuer une analyse hiérarchique de scène en se reposant sur deux processus : ascendant et descendant. Le processus ascendant permet d'accumuler l'évidence sur l'existence des régions (ou objets), tandis que le processus descendant permet de mettre en cause, par l'expert, le contenu informationnel des régions et objets identifiés dans le processus ascendant. En effet, ce système met en place les outils possibilistes nécessaires à l'exploitation conjointe de plusieurs sources de connaissances (fusion) afin d'améliorer l'interprétation des données acquises en vue d'une représentation plus riche de la scène observée : les deux principales applications visées sont donc la segmentation / classification (processus ascendant) et le démixage (processus descendant).

Les performances du système d'interprétation proposé sont évaluées en utilisant plusieurs types d'images (images de synthèse, images mammographiques, images du type IRM, images SONAR, et Images satellitaires). Les résultats obtenus sont très encourageants et montrent l'efficacité de la théorie des possibilités comme un cadre de représentation des connaissances a priori et comme outil de raisonnement permettant d'extraire de nouvelles connaissances à partir des connaissances disponibles. De plus, l'approche proposée s'est montrée très efficace pour l'intégration de plusieurs sources des connaissances.

Mots-clés : Interprétation de scène, Théorie de possibilités, Raffinement de distribution de possibilités, Mesure de similarité, Fusion de connaissances.

Abstract

This thesis considers the context of the possibilistic knowledge representation and processing in order to develop a scene interpretation system observed by multiple sensors. The research is focusing on a possibilistic knowledge modeling of the imperfect information provided by the sensors. Indeed, this formalism allows introducing and explicitly manipulating, within the same framework, certain imperfections carried by the collected information and also some priori knowledge about the scene, sensors, and acquisition conditions expressed by experts or resulting from observations or ambiguous judgments.

This work is divided into several stages: study of the state of the art and mastery of the possibility theory, hierarchical/semantic scene analysis, study of the application of possibility theory in different semantic levels (sub-pixel, pixel, region, object and scene), and finally the application of the proposed approach for scene interpretation.

The final objective of this thesis is to develop a scene interpretation system based on a possibilistic approach. This approach performs a hierarchical scene analysis relying on two processes: ascending and descending processes. The ascending process allows accumulating evidences on the existence regions (or objects), while the descending process allows calling into question, by the expert, the informational content of the regions and objects identified in the ascending process. Indeed, this system sets up the possibilistic tools allowing to jointly exploiting multiple sources of knowledge (fusion) in order to improve the interpretation of the acquired data and, thus, enriching the representation of the observed scene. The two main performed applications are segmentation/classification (ascending process) and the unmixing (descending process).

The performance of the proposed system is evaluated using several types of images (synthetic images, mammographic images, MRI images, SONAR images and satellite images). The obtained results are very encouraging and show the effectiveness of possibility theory as a framework for priori knowledge representation and as a reasoning tool to extract new knowledge from available knowledge. Moreover, the proposed approach has been very effective for integrating multiple knowledge sources.

Keywords : Scene interpretation, Possibility theory, Refining of possibility distribution, Similarity measure, Knowledge fusion.